



```

XX Sequence 573 AA;
SQ
Query Match
Best Local Similarity 100.0%; Score 2910; DB 3; Length 573;
Matches 573; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MITALDLVHLTAIVVPLVAVMTLAYGSRVWRWRIPTDQCSCGINRFVALFAVPLLSFHFIS 60
DB 1 MITALDLVHLTAIVVPLVAVMTLAYGSRVWRWRIPTDQCSCGINRFVALFAVPLLSFHFIS 60

QY 61 TNDPFAMNRLFLAATLOKVAIVLALLALASGLSSPRALGLDWSITLFSITLPTLVWG 120
DB 61 TNDPFAMNRLFLAATLOKVAIVLALLALASGLSSPRALGLDWSITLFSITLPTLVWG 120

QY 121 IPLLRLGMYGASAGTLMVQVVLQCIITWYTLMLFLFYRAARALVLDQPDGAAASIVSF 180
DB 121 IPLLRLGMYGASAGTLMVQVVLQCIITWYTLMLFLFYRAARALVLDQPDGAAASIVSF 180

QY 181 RVDSDVLSLARGDVELEAPDGVAGAVSRGGDAGRVRVTVRKSTSSRSEAAASHSHS 240
DB 181 RVDSDVLSLARGDVELEAPDGVAGAVSRGGDAGRVRVTVRKSTSSRSEAAASHSHS 240

QY 241 QTMQPRVNLSGVEIYLSQSSRNPTPRGSSFNHADFNIVGAAGKGGGAAGDEEKGACG 300
DB 241 QTMQPRVNLSGVEIYLSQSSRNPTPRGSSFNHADFNIVGAAGKGGGAAGDEEKGACG 300

QY 301 GGGGSHSPQPOAVAPAKRDLMLVWSSSASPVSERAAVHVFGAGADHADVLAKGAQA 360
DB 301 GGGGSHSPQPOAVAPAKRDLMLVWSSSASPVSERAAVHVFGAGADHADVLAKGAQA 360

QY 361 YDEYGRDYSRTKNGSGGADKGGPTLSKLGSNSTAQLYPKDDGEGRAAAMPASVMT 420
DB 361 YDEYGRDYSRTKNGSGGADKGGPTLSKLGSNSTAQLYPKDDGEGRAAAMPASVMT 420

QY 421 RLILIMWRKLRNPNTYSSLIGVWVSLVSRWGIEMPAIIRSIISLSDAGLGMAMFSL 480
DB 421 RLILIMWRKLRNPNTYSSLIGVWVSLVSRWGIEMPAIIRSIISLSDAGLGMAMFSL 480

QY 481 GLFWALQRIIACGNKLAAMGVRFVAGPAVMAAASIAVLGVLHLLHAIIVQALPOQI 540
DB 481 GLFWALQRIIACGNKLAAMGVRFVAGPAVMAAASIAVLGVLHLLHAIIVQALPOQI 540

QY 541 VPFFVFAKEYGVHPDILSTAYGPIITSHGFITCHS 573
DB 541 VPFFVFAKEYGVHPDILSTAYGPIITSHGFITCHS 573

RESULT 2
RAY44265
ID AAY44265 standard; protein; 595 AA.
AC AAY44265;
XX
DT 28-FEB-2000 (first entry)
DE Rice EIR1 Homologue (REH1) protein.
KW Root-specific auxin transport protein; EIR1; Ethylene Insensitive Root;
KW gravitropism; auxin-based herbicide; plant root; transgenic plant;
KW auxin homeostasis; Rice EIR1 Homologue; REH1; transmembrane domain.
XX
OS Oryza sativa.
XX
FH Key
FH Peptide
FT Location/Qualifiers
FT 1..60
FT /label= Signal_peptide
FT 10..28
FT /label= Transmembrane_domain
FT 41..60
FT /label= Transmembrane_domain
FT 71..92
FT /label= Transmembrane_domain

```

```

FT Domain 100..121
FT /label= Transmembrane_domain
FT Domain 132..152
FT /label= Transmembrane_domain
FT Domain 451..468
FT /label= Transmembrane_domain
FT Domain 478..501
FT /label= Transmembrane_domain
FT Domain 516..537
FT /label= Transmembrane_domain
FT Domain 543..559
FT /label= Transmembrane_domain
FT Domain 570..590
FT /label= Transmembrane_domain

XX WO9963092-A1.
XX 09-DEC-1999.
XX 03-JUN-1999; 99WO-US012277.
XX 03-JUN-1998; 98US-0087789P.
XX (MHED) WHITEHEAD INST BIOMEDICAL RES.
XX Luschnig C, Gaxiola RA, Grisafi P, Fink GR;
XX WPI; 2000-086979/07.
XX N-PSDB; AAZ29280.
XX DNA encoding a root-specific auxin transport protein, used to develop
XX transgenic plants with increased resistance to herbicides.
XX Claim 18; Fig 9; 55pp; English.

The present sequence is a REH1 protein, a rice homologue of EIR1 which
functions as a root-specific auxin transport (efflux) protein involved in
gravitropism. The sequence is obtained from a rice EST derived from root-
specific cDNA. REH1 consists of potential N-glycosylation sites and a N-
terminal signal peptide. The transmembrane domains are located in the
highly conserved portions of the protein. N- and C-termini. The sequence
is used for producing genetically engineered plants with greater
resistance to auxin-based herbicides and auxin transport inhibitors in
combination with a second herbicide. It can also be used to enhance
transport of auxin in plant roots, produce transgenic plants which
exhibit altered auxin homeostasis and mutant plants in which the roots
are agravitropic and have reduced sensitivity to ethylene

XX Sequence 595 AA;
SQ
Query Match 57.2%; Score 1665; DB 3; Length 595;
Best Local Similarity 58.9%; Pred. No. 3e-147;
Matches 365; Conservative 54; Mismatches 111; Indels 90; Gaps 17;

QY 1 MITALDLVHLTAIVVPLVAVMTLAYGSRVWRWRIPTDQCSCGINRFVALFAVPLLSFHFIS 60
DB 1 MITADDFYHVTAMVPLVAVMTLAYGSRVWRWRIPTDQCSCGINRFVALFAVPLLSFHFIS 60

QY 61 TNDPFAMNRLFLAATLOKVAIVLALLA---LASGLSSPRALGLDWSITLFSITLPT 116
DB 61 TNNPYTNLRFIAATLOKVAIVLALLA---SLEWITLFSITLPT 113

QY 117 LVNMGIPLLRGMYGASAGTLMVQVVLQCIITWYTLMLFLFYRAARALVLDQPDGAAAS 176
DB 114 LVNMGIPLLRGMYGFS--GSLMVQIVVLQCIITWYTLMLFMFYRGARMLITEQFPD--TAAN 171

QY 177 IVSFRVDSDVLSL--ARGDVELEAPDGVAGAVSRGGDAGRVRVTVRKSTSSRSEA 234
DB 172 IASIVVDVDSVLSLGGRRDAITETEVK-----EDGRHVTVRKSNASRSIIY 218

QY 235 CSHSHS-QTMQPRVNLSGVEIYLSQSSRNPTPRGSSFNHADFNIVGAAGKGGGAAGD 293
DB 219 SRSMGFSSTTPRPSNLTNAEYLSQSSRNPTPRGSSFNHADFNIVGAAGKGGGAAGD 278

```



294	EEKGAC-----GGGGGH--SPQQAQAVAP-----AKR	319
Db		
279	VETGATPSPSYEDDASPKPKPLPASNAAPWAGHYPAFNPFAVSSAPGAKKAATNGQAKG	338
Qy		
320	KDLHMLVWSSSSASPVSERAAAVHVFGAGGADHADVLA-----KGAQAYDEY-GRDQYSS	371
Db		
339	EDLHMFVWSSSSASPSVD-----VFGGAPDYNDAAAVKSPRKMDGAKREDYVDERDDFSF	393
Qy		
372	RTKNGSGGADGGGPTLSKLGNSSTAQIYPPKDDGEGRAAAVAMPPPASVMTLILIMVWRKL	431
Db		
394	GNR-----GYMDRD-----AEAGDEKAAAAAGADPPKAMAAPTAMPTPTSVMTLILIMVWRKL	446
Qy		
432	IRNPNTYSLLIGVWMSLVSYPWGTEMPAIIARTSILSDAGLGWAFPSGLFMALQPHII	491
Db		
447	IRNPNTYSLLIGLINSLVCFRNFEMPAIVLKSILSDAGLGWAFPSGLFMALQPHII	506
Qy		
492	ACGNKLAIAIMGVRFPAGPAPMAAASITAVGLRGVLLHIAIVQAALPQGIYVPVFPAKEYGV	551
Db		
507	ACGNKVATYAMAVRFLAGPAPMAAASFAVLGRGTLHVAIVQAALPQGIYVPVFPAKEYSV	566
Qy		
552	HPDILSTA--YG-----GIT	564
Db		
567	HPSILSTAVIFGMLIALPIT	586

RESULT 3	
ABB91576	
ID	ABB91576 standard; protein; 622 AA.
XX	
AC	ABB91576;
XX	
XX	31-MAY-2002 (first entry)
XX	
DE	Herbicideally active polypeptide SEQ ID NO 787.
XX	
KW	Herbicideal; plant; agriculture; herbicide.
XX	
OS	Arabidopsis thaliana.
XX	
PN	WO200210210-A2.
XX	
PD	07-FEB-2002.
XX	
PF	28-AUG-2001; 2001WO-EP009892.
XX	
PR	28-AUG-2001; 2001WO-EP009892.
XX	
PA	(FARB ) BAYER AG.
XX	
PI	Tietjen K, Weidler M;
XX	
DR	WPI; 2002-269010/31.

RESULT 4	
AAB26942	
ID AAB26942 standard; protein; 60S AA.	
XX XX AC AC	
AA B26942;	
XX XX DT DT	(first entry)
DE DE	Soybean auxin transport protein clone sfl.p.k131.gg.
KW KW	Auxin transport protein; soybean; root development; gene mapping;
XX KW	plant breeding; herbicide.
Glycine max.	
OS OS	
P N PN	WO2000068389-A2.
PD PD	16-NOV-2000.
PF PF	03-MAY-2000; 2000WO-US012061.
PR PR	07-MAY-1999; 99US-0133040P.
PA PA	(DUPO ) DU PONT DE NEMOURS & CO E I .

PA (PION-) PIONEER HI-BRED INT INC.  
 XX Orozco EM, Weng Z, Bruce WB, Cahoon RE, Tao Y;  
 XX WPI; 2000-687647/67.  
 DR N-PSDB; AAA94729.  
 XX New nucleic acid sequences encoding new auxin transport proteins, useful  
 PT for modulating root growth of plants and to screen for herbicides.  
 XX Claim 10; Fig 1; 94pp; English.  
 XX Auxins are plant hormones that influence plant behaviour and development  
 CC e.g. vascular tissue differentiation, apical development, tropic  
 CC responses and organ (e.g. flower, leaf) development. The present  
 CC invention relates to auxin transport proteins and coding sequences. The  
 CC present sequence is one such auxin transport protein. The DNA sequence  
 CC encoding this protein may be used to modulate root development, e.g. to  
 CC produce a more robust root system, alter root angle or redirect root  
 CC growth; and may be useful for gene mapping (e.g. for plant breeding) and  
 CC to identify loss of function mutants. This protein may be useful for  
 CC raising specific antibodies, for the detection of auxin transport  
 CC proteins and to design and/or identify specific inhibitors of auxin  
 CC transport proteins, potentially useful as herbicides  
 XX Sequence 605 AA;  
 SQ  
 Query Match 56.5%; Score 1643; DB 3; Length 605;  
 Best Local Similarity 57.3%; Pred. No. 3.6e-145;  
 Matches 361; Conservative 58; Mismatches 111; Indels 100; Gaps 17;  
 QY 1 MITALDLVHLTAVVPLVYAMTLAGSVRWRIPTDQCSCGINRFVALFAVPLLSFHFTS 60  
 DB 1 MITLDFHVTAMVPLVYAMTLAGSVRWKIFSPDQCSCGINRFVALFAVPLLSFHFTS 60  
 QY 61 TNDPFAMNLFRAADTLQKVAVLALALASRGLSPRALGLDWSITLPSLSTPLNTLYMG 116  
 DB 61 SNPYENMLFLAADTLQKIILVLLAVMSNITKRG-----LEWATLPSLSTPLNT 113  
 QY 117 LVMGTPILLRGMYGASSAGTLMVQVVLQCIITWTMLFLPEYRAARALVLDOPDGAAS 176  
 DB 114 LVNGIPLKMGYDPS-GSLMVQIVLQCIITWTMLFLFEFGARMLISEQPD-TAAS 171  
 QY 177 IVSFVSDVSVSL-ARGDVELEAPDPGVAGAGVSSRGGDAGRVVTVRKSTSSRSEAC 235  
 DB 172 IVSIVHSDVMSLDGROPLETEAIK-----EDGKLHVTVRKSNASRSDIFS 218  
 QY 236 SHSHS-QTMQPRVSNLGSVEIYSQSSRNPTPRGSSFNHADFFNIYGAAKGGGGAAGD- 293  
 DB 219 RSQGLSTTPRPNLNTVAEIIYSQSSRNPTPRGSSFNHDTFYSMAAGRSNFCASDV 278  
 QY 294 -----EKG-----CGGGGGHSPQ-----AVAPAK 318  
 DB 279 YGLSASRGFTPRPSNYDEGKPFHYHAAGTGTHYPAPNPGMFSPSGSKSYAANNAK 338  
 QY 319 R-----KDLHMLWSSSSASPVSEARAHHVFGAGGADHADVLAKGAQAYDEYG 365  
 DB 339 RPNQAQLKPDGHRDLHMFWSSSASPVSDVFGAHEYG-CGHQKEVKLVNSPGKVENN 397  
 QY 366 ---RDDYSSTKXSGGADGGPTLSKLSNSTAQLYPKDDGEGRAAAVAMPVPASVMT 421  
 DB 398 HRDTQEDYLEKDEFSGNREM-----DRENNQLEGEKVGDKPK--TWPPASVMT 446  
 QY 422 LILNWKRLRNPNNTYSSLLGVVSVSVKRWGTEPAIARSISILSDAGLGMAMFSLG 481  
 DB 447 LILNWKRLRNPNNTYSSLLGLTWSLVSPKRVNEMPAIARSISILSDAGLGMAMFSLG 506  
 QY 482 LFMALQPRIIACGNKLAANGVRFVAGPAMAAASIAVGLRGVLLHIAIVQALPQGV 541  
 DB 507 LFMALQPRVIACGNSTAFANAVRFLTGPAMAAASIAVGLKGVLLHFAIVQALPQGV 566  
 QY 542 PVPFAKEYGVHDPDILSTA--YG-----PIT 564

DB 567 PVPFAKEYGVHDPDILSTAVIFGMLIALPIT 596  
 RESULT 5  
 AAB26950  
 ID AAB26950 standard; protein; 622 AA.  
 XX AAB26950;  
 AC AAB26950;  
 XX 02-FEB-2001 (first entry)  
 DE Auxin transport protein AtPIN1.  
 KW Auxin transport protein; corn; root development; gene mapping;  
 XX plant breeding; herbicide.  
 OS Arabidopsis thaliana.  
 XX WO200068389-A2.  
 XX 16-NOV-2000.  
 XX 03-MAY-2000; 2000WO-US012061.  
 XX 07-MAY-1999; 99US-0133040P.  
 PA (DUPO) DU PONT DE NEMOURS & CO E I.  
 PA (PION-) PIONEER HI-BRED INT INC.  
 XX Orozco EM, Weng Z, Bruce WB, Cahoon RE, Tao Y;  
 XX WPI; 2000-687647/67.  
 PT New nucleic acid sequences encoding new auxin transport proteins, useful  
 PT for modulating root growth of plants and to screen for herbicides.  
 XX Disclosure; Fig 1; 94pp; English.  
 CC Auxins are plant hormones that influence plant behaviour and development  
 CC e.g. vascular tissue differentiation, apical development, tropic  
 CC responses and organ (e.g. flower, leaf) development. The present  
 CC invention relates to auxin transport proteins and coding sequences. The  
 CC present sequence is one such auxin transport protein. The DNA sequence  
 CC encoding this protein may be used to modulate root development, e.g. to  
 CC produce a more robust root system, alter root angle or redirect root  
 CC growth; and may be useful for gene mapping (e.g. for plant breeding) and  
 CC to identify loss of function mutants. This protein may be useful for  
 CC raising specific antibodies, for the detection of auxin transport  
 CC proteins and to design and/or identify specific inhibitors of auxin  
 CC transport proteins, potentially useful as herbicides  
 XX Sequence 622 AA;  
 SQ  
 Query Match 56.4%; Score 1641.5; DB 3; Length 622;  
 Best Local Similarity 56.1%; Pred. No. 5.2e-145;  
 Matches 365; Conservative 56; Mismatches 105; Indels 125; Gaps 18;  
 QY 1 MITALDLVHLTAVVPLVYAMTLAGSVRWRIPTDQCSCGINRFVALFAVPLLSFHFTS 60  
 DB 1 MITAADFVHTAMVPLVYAMTLAGSVRWKIFTPDQCSCGINRFVALFAVPLLSFHFTS 60  
 QY 61 TNDPFAMNLFRAADTLQKVAVLALALASRGLSPRALGLDWSITLPSLSTPLNTLYMG 120  
 DB 61 ANNPYAMNLFRAADSLQKIVLVLFLWCK-LS-RNGSLDWTITLPSLSTPLNTLYMG 117  
 QY 121 IPLLGRMGYSSAGTLMVQVVLQCIITWTMLFLPEYRAARALVLDOPDGAASIVSF 180  
 DB 118 IPLLGRMGY-NFGDGLMVQIVLQCIITWTMLFLPEYRAKLLISEQPD-TAGSIYSI 175  
 QY 181 RVSDVSVSL-ARGDVELEAPDPGVAGAGVSSRGGDAGRVVTVRKSTSSRSEACSHH 239  
 DB 176 HVSDIMSLDGRQPLETEAIK-----EDGKLHVTVRKSNASRSDIYSRSRQ 222



XX PD 06-SEP-2000.  
XX PF 25-FEB-2000; 2000EP-00301439.  
XX PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 05-MAY-1999; 99US-0132484P.  
PR 06-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 11-MAY-1999; 99US-0132863P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135623P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 28-MAY-1999; 99US-0136392P.  
PR 01-JUN-1999; 99US-0136782P.  
PR 03-JUN-1999; 99US-0137222P.  
PR 04-JUN-1999; 99US-0137528P.  
PR 07-JUN-1999; 99US-0137502P.  
PR 08-JUN-1999; 99US-0137724P.  
PR 10-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 14-JUN-1999; 99US-0138847P.  
PR 16-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 22-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 28-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.

PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 18-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144864P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147102P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 25-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.

PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161933P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 53.6%; Score 1560.5; DB 3; Length 640;  
Best Local Similarity 52.9%; Pred. No. 2.2e-137;  
Matches 353; Conservative 73; Mismatches 102; Indels 139; Gaps 20;

QY 1 MITALDLHYLTAVPLVYAMTILAYGVSRWRIPTDQCSGINRFVAFVPLLSFHFIS 60  
DB 1 MISHWDLTYVLTAVPLVYAMTILAYGVSRWRIPTDQCSGINRFVAFVPLLSFHFIS 60  
QY 61 TNDPFAMNLFRAADTLQKAVLALALASRLGLSDWITLPSLTPLNTLVWG 120  
DB 61 TNNPYAMNLFRAADTLQKIMLSLVLWA--NFTSGSLEWSTIFSLSTPLNTLVWG 117  
QY 121 IPLLRCWYGASSAGTLMVQVVLQCTIIVTLMLEFVYRAAEALVLDOPPDGAASIVSF 180  
DB 118 IPLLRYMYEYS-GSLMVQIVLQCTIIVTLMLEFVYRAAEALVLDOPPDGAASIVSF 175  
QY 181 RVDSVDVVLARGD-VELEAPDPGVAGAGAVSRGGDAGRVRVTVKRTSSRSEACSHSH 239  
DB 176 KVESDVVSLDGHDFLETDAEI-----GDDGKLHVTVKSNASR-RSFC--- 217  
QY 240 SQTMQVRVNLQSVETYSIQSRNPTPGSSFHEDAFNIIVG-----A 282  
DB 218 GPNWTPRPNLTGAELYSJST---TPRGSNHNHSDFTNMWGFPGRLSNFGPADMYSVQ 273  
QY 283 AAKG-----GGGAAGDEKGAACGGCGGHSPQP----- 310  
DB 274 SSRGTPRPSNFENECAMASSPRFGYPPGGGAGSYAPAPNPFSSTTTSTANKSVNKNPKD 333  
QY 311 ----QAVAVPA-----KXKDLHMLVWSSASPVYERAAVHVFAGGAGADR-----ADVL 354  
DB 334 VNTNQTTTLPTGKSNSHDAKELHFWVSSNGSPVSDRAGLNVFG-GAPDNDQGRSDQG 392  
QY 355 AK-----GAQAYDEYGRDDYSRTRKNGSGGADKGGPTL 387

DB 393 AKETRLVDPQSHNGETKAVAHPASGDGFGCEQOFQFAGKEEAEERPKDAENG-----L 445  
QY 388 SKLGSNSTAQLYPKDDEGCRRAAVA---MPPASVNTRLILIMVWKLIRNPNVYSSLIGV 444  
DB 446 NKLAPNSTAALQSK-TGLGGRASQRKMPASVNTRLILIMVWKLIRNPNVYSSLIGL 504  
QY 445 VWSLVYRMGIEMPAIIARSISILSDAGLGMAWFSGLFMALOPRIIACGNKLAIAINMGV 504  
DB 505 IVALVAFRWHVAMPKTIQOSISILSDAGLGMAWFSGLFMALOPKLIACGNSVATFAMAV 564  
QY 505 RFVAGPAVMAAASIAVGLRGVLLHAIQAAALPOQGIIVPFAKEVGVHPDITLSTA--YG- 561  
DB 565 RFTGPAVMAVAIAIGLRGDLRLVAIVQAAALPOQGIIVPFAKEVGVHPDITLSTA--YG- 624  
QY 562 ----PIT 564  
DB 625 LIALPIT 631

RESULT 8  
ABB91542  
ID ABB91542 standard; protein; 640 AA.

XX ABB91542;  
XX 31-MAY-2002 (first entry)  
XX Herbicidally active polypeptide SEQ ID NO 753.  
XX Herbicidal; plant; agriculture; herbicide.  
XX Arabidopsis thaliana.  
XX WO200210210-A2.  
XX 07-FEB-2002.  
XX 28-AUG-2001; 2001WO-EP009892.  
XX 28-AUG-2001; 2001WO-EP009892.  
XX (FARB ) BAYER AG.  
XX Tietjen K, Weidler M;  
XX WPI; 2002-269010/31.  
XX

Identifying plant target proteins for herbicidally active compounds,  
comprising aligning and comparing nucleic acid or amino acid sequences  
from plant with nucleic acid or amino acid sequences from non-plant  
organisms.

Claim 5; SEQ ID NO 753; 261pp + Sequence Listing; English.

The invention relates to identifying target proteins (ABB90790-ABB94016)  
for herbicidally active compounds, comprising aligning and comparing  
nucleic acid or amino acid sequences from plant with nucleic acid or  
amino acid sequences from non-plant organisms using suitable search  
parameters, where plant sequences having an E-value greater by a factor  
of 3 than the E-value of most similar non-plant sequences are selected.  
The polypeptides or nucleic acids encoding them are useful for  
identifying modulators. The identified modulators are useful as  
herbicides

Sequence 640 AA;

Query Match 53.6%; Score 1560.5; DB 5; Length 640;  
Best Local Similarity 52.9%; Pred. No. 2.2e-137;  
Matches 353; Conservative 73; Mismatches 102; Indels 139; Gaps 20;  
QY 1 MITALDLHYLTAVPLVYAMTILAYGVSRWRIPTDQCSGINRFVAFVPLLSFHFIS 60

Db 1 MTSWHDLYTLVTAVIPLYVAMILAYGVSRVKKIFSPQDCSGINRPFVAIFAVPILLSFHFIS 60  
QY 61 TNDPFAANRFLAADTLQKAVVALLALASRGLSPRALGLDWSITLESITLNTLYMG 120  
Db 61 TNNPYANRFLAADTLQKIMLSILVLA---NTRSGSLEWSTIFSLSTLNTLYMG 117  
QY 121 IPLLKMYGASSAGTLMVQVWVLCIIWTLMLFPEYRAARALVLDQFPDGAASIVSF 180  
Db 118 IPELLAMYGEYS-GSLMVOIVVLCIIWTLMLFPEYRAARALVLDQFPDGAASIVSF 175  
QY 181 RVSDVVSILARGD-VELEAPDGVAGAGAVSRGGDAGRVTYKSTSSSEACSHSH 239  
Db 176 KVESDVVSLDGHDFLETDAEI-----GDDGKLHVTVRKSNAASR-RSFC----- 217  
QY 240 SQTMOPRVNLGSVEIYLSQSSRNPTPRGSSFNHADFFNIVG-----A 282  
Db 218 GNMTPRPSNLTCAGIYLSLT-----TPRGSNFNHSDFYNMGFPDGLSNGPADMYSVQ 273  
QY 283 AAKG-----GGGAAGDEKACGGGGGHHSPQ----- 310  
Db 274 SSRGTPRPSNFEENCAMASSPRFGYPGGGAGSYPAAPNPEFSSTTTSTANKSVNKNPKD 333  
QY 311 ----QAVAVPA-----KKDLHMLVWSSASPYSERAAVHVFGAGGADH-----ADVL 354  
Db 334 VNTNQOTTLTGTGKSNHDAKELHMFVWSNGSPVSDRAGLVFG-GAPDNDGGRSQDQ 392  
QY 355 AK-----GAQAYDEYGRDYSSTKXGSGGADKGPTL 387  
Db 393 AKRMLVPDQSHNGETKVAHPASGDFGGEQOFSPAGKEEAERPQDAENG-----L 445  
QY 388 SKLGSNSTACLYKDDGEGRAAVA---MPPASVMTLRLILMVWRKLIENPTYSSLIGV 444  
Db 446 NKLAPNSTAALQSK-TGLGAEASQRKNMPPASVMTLRLILMVWRKLIENPTYSSLIGL 504  
QY 445 VMSIVSRVWGIEMPAIARSISILSDAGLGMAMFSLGLFMALQPRIIACNKLAAIAMGV 504  
Db 505 INALVAFRWHVAMPKIIQQSISILSDAGLGMAMFSLGLFMALQPKLIACGNSVATPAMAV 564  
QY 505 RFVAGPAMVMAASIANVGLRVLHIALVQALPQGIVPVFVFAKEYGVHPDILSTA-YG- 561  
Db 565 RFLTGPAVMAVAATAIGRLGDLRVAIVQALPQGIVPVFVFAKEYNVHPAILSTGVIFGM 624  
QY 562 ----PIT 564  
Db 625 LIALPIT 631

RESULT 9

AAG31983

ID AAG31983 standard; protein; 616 AA.

AC AAG31983;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 38502.

KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; Genetic mapping; gene expression control; promoter;  
KW termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-00301439.

PR 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129645P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 06-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136382P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 07-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139432P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144003P.  
PR 16-JUL-1999; 99US-0144085P.

PR	16-JUL-1999;	99US-0144086P.	PR	08-OCT-1999;	99US-0158232P.
PR	19-JUL-1999;	99US-0144325P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0158293P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159329P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159637P.
PR	20-JUL-1999;	99US-0144684P.	PR	14-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-0144814P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160767P.
PR	21-JUL-1999;	99US-0145085P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.
PR	26-JUL-1999;	99US-0145226P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161406P.
PR	27-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161920P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161922P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161933P.
PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.
PR	04-AUG-1999;	99US-0147302P.			
PR	05-AUG-1999;	99US-0147192P.			
PR	05-AUG-1999;	99US-0147260P.			
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.			
PR	09-AUG-1999;	99US-0147493P.			
PR	09-AUG-1999;	99US-0147935P.			
PR	10-AUG-1999;	99US-0148171P.			
PR	11-AUG-1999;	99US-0148319P.			
PR	12-AUG-1999;	99US-0148341P.			
PR	13-AUG-1999;	99US-0148565P.			
PR	13-AUG-1999;	99US-0148684P.			
PR	16-AUG-1999;	99US-0149368P.			
PR	17-AUG-1999;	99US-0149175P.			
PR	18-AUG-1999;	99US-0149428P.			
PR	20-AUG-1999;	99US-0149722P.			
PR	20-AUG-1999;	99US-0149723P.			
PR	23-AUG-1999;	99US-0149902P.			
PR	23-AUG-1999;	99US-0149929P.			
PR	23-AUG-1999;	99US-0149930P.			
PR	25-AUG-1999;	99US-0150566P.			
PR	26-AUG-1999;	99US-0150884P.			
PR	27-AUG-1999;	99US-0151065P.			
PR	27-AUG-1999;	99US-0151066P.			
PR	27-AUG-1999;	99US-0151080P.			
PR	30-AUG-1999;	99US-0151303P.			
PR	31-AUG-1999;	99US-0151438P.			
PR	01-SEP-1999;	99US-0151930P.			
PR	07-SEP-1999;	99US-0152363P.			
PR	10-SEP-1999;	99US-0153070P.			
PR	13-SEP-1999;	99US-0153758P.			
PR	15-SEP-1999;	99US-0154018P.			
PR	16-SEP-1999;	99US-0154039P.			
PR	20-SEP-1999;	99US-0154779P.			
PR	22-SEP-1999;	99US-0155139P.			
PR	23-SEP-1999;	99US-0155486P.			
PR	24-SEP-1999;	99US-0155659P.			
PR	28-SEP-1999;	99US-0156458P.			
PR	29-SEP-1999;	99US-0156596P.			
PR	04-OCT-1999;	99US-0157117P.			
PR	05-OCT-1999;	99US-0157753P.			
PR	06-OCT-1999;	99US-0157863P.			
PR	07-OCT-1999;	99US-0158029P.			
Query Match 53.0%; Score 1542.5; DB 3; Length 616;					
Best Local Similarity 53.5%; Pred. No. 1e-135;					
Matches 349; Conservative 59; Mismatches 111; Indels 133; Gaps 17;					
Qy	1	MITALDLXVLTAVVPLVYVAMTILAYGVSVMWRIFTPDQCSGINRFFVALFAVPLLSPHFIS 60			
Db	1	MITWHDLYVLTAVVPLVYVAMTILAYGVSVMWRIFTPDQCSGINRFFVALFAVPLLSPHFIS 60			
Qy	61	TNDPFAMNLFRAADTLQKVALALALASRGSLSPRALGLDWSITLPSLSTLPTLVMG 120			
Db	61	TNDPFAMNLFRAADTLQKIMLVLLALWA---NLTKGSLWMTITFSLSTLPTLVMG 117			
Qy	121	IPLLRGWYGASSAGTLMVQVVLQCIITWTMLFLFEYRAARALVLDQPDGAAASIVSF 180			
Db	118	IPLLIANYG-TYAGSLMVQVVLQCIITWTMLFLFEYRAARALVLDQPDGAAASIVSF 175			
Qy	181	RVDSVYVSLARGD-VELEAEPDVGAGAVSRGGDAGVRVTVRKSTSRSEACSHSH 239			
Db	176	KVESDVVSLDGHDFLEDAEI-----GNDGKLHVTVRKSNASR-----R 214			
Qy	240	SQTMQPRVNLQSVIYISLQSSNPPTPGSSFNHADFFNIVGAAKGGGAAGD----- 293			
Db	215	SLMTPRPNLTGAEIYSLSS-----TPRGSNFNHSDFYSVMGFP-----GRLSNFQPADL 266			
Qy	294	-----EEKGACGGG--GGHSPQPOAVAPAKR----- 319			
Db	267	YSVQSSRGPTPRPSNFENNAVKYGFYNNNTSSVPAAGSVAPAPNPEFTGTGVSTKPNKI 326			
Qy	320	-----KDLHMLVWSSASPVSERAAHVHVGAGADHADVLAKGAQA 360			
Db	327	PKENQQLOEKDKSKASHDAKELHMFVWSSASPVSD-----VFGGAGDNV-----ATE 375			
Qy	361	YDEYGRDDY-----SSRTKNGSGGAD-----KGGPTLKLKLSNSTAQLY 399			
Db	376	QSEQGAKEIRMVVSDQPRKSNARGGGDDIGGLDSGGERIEKATAGLNKMGNSSTAELE 435			
Qy	400	PKDDGEGRAAAVAMPASVMTLILIMVWRKLIRNPTVSSLLIGVWLSVYRWIEMPA 459			
Db	436	AAGDGGGNGNTHMPPTSVMTLILIMVWRKLIRNPTVSSLLIGLWALVYVHWVAMPK 495			
Qy	460	IIARSTISLSDAGLGMFSLGLFMALQPRIIACGNKLAIAAMGVRFVAGPAAVMAAASIA 519			



Db 496 ILQOSISILSDAGLGMAMFSLGLFVALQPKIIACGNSVATFAMAVRFTITGPAIMAVAGIA 555  
 QY 520 VGLRGVLLHIAIVQAAALPQGIIVFVFAKEYGVHPDILSTA--YG-----PIT 564  
 Db 556 IGLHGDLLRIAIVQAAALPQGIIVFVFAKEYGVHPDILSTA--YG-----PIT 607

RESULT 10  
 ID ABB91685 standard; protein; 616 AA.  
 AC ABB91685;  
 XX  
 DT 31-MAY-2002 (first entry)  
 DE Herbicidally active polypeptide SEQ ID NO 896.  
 DE Herbicidal; plant; agriculture; herbicide.  
 KW Arabidopsis thaliana.  
 OS Arabidopsis thaliana.  
 XX  
 FN WC200210210-A2.  
 PD  
 PD 07-FEB-2002.  
 XX  
 XX 28-AUG-2001; 2001WO-BP009892.  
 XX  
 XX 28-AUG-2001; 2001WO-BP009892.  
 PR (FARB ) BAYER AG.  
 PA Tietjen K, Weidler M;  
 PI WPI; 2002-269010/31.  
 DR  
 XX Identifying plant target proteins for herbicidally active compounds,  
 PT comprising aligning and comparing nucleic acid or amino acid sequences  
 PT from plant with nucleic acid or amino acid sequences from non-plant  
 PT organisms.  
 XX  
 PS Claim 5; SEQ ID NO 896; 261pp + Sequence Listing; English.  
 XX  
 XX The invention relates to identifying target proteins (ABB90790-ABB94016)  
 CC for herbicidally active compounds, comprising aligning and comparing  
 CC nucleic acid or amino acid sequences from plant with nucleic acid or  
 CC amino acid sequences from non-plant organisms using suitable search  
 CC parameters, where plant sequences having an E-value greater by a factor  
 CC of 3 than the E-value of most similar non-plant sequences are selected.  
 CC The polypeptides or nucleic acids encoding them are useful for  
 CC identifying modulators. The identified modulators are useful as  
 CC herbicides  
 XX  
 SQ Sequence 616 AA;

Query Match 53.0%; Score 1542.5; DB 5; Length 616;  
 Best Local Similarity 53.5%; Pred. No. 1e-135;  
 Matches 349; Conservative 59; Mismatches 111; Indels 133; Gaps 17;

QY 1 MITALDLHYLTAVPLVYAMTAYSGVRWRIFTDQCSGINRFVAFVLLSFHFTS 60  
 Db 1 MITWHDLYTLTAVPLVYAMTAYSGVRWRIFTDQCSGINRFVAFVLLSFHFTS 60  
 QY 61 TNDPPAMNLFRLAADTLQKAVLALLASRGLSSPRALGLDWSITLFLSLTLPNTLVNG 120  
 Db 61 TNDPPAMNFRFAADTLQKIMVLLALWA---NLTKNGSLEWMTIFLSLTLPNTLVNG 117  
 QY 121 IPLLIRMGYASSATLMMQVVLQCIITWYTLMLFLFEYRAARALVLDQPDGAASIVSF 180  
 Db 118 IPLLIRMGY-TYAGSLMVQVVLQCIITWYTLMLFLFEYRGAKLLIMEQFPE-TGASIVSF 175  
 QY 181 RVSDVSVLSLARGD-VELEAEFDGVAGAVSRGGDAGRVTVTRKSTSRSEAAASHH 239  
 Db 176 KVESDVVSLDGHDFLETDAE:-----GNDGKLHVTVRKSNASR-----R 214

QY 240 SQTMRVSNLQVEIYSLQSRNPTPRGSSFNHADFFNIVGAAAKGGGGAAGD----- 293  
 Db 215 SLMTFRPNLTGAEIYSLSS-----TPRGSNFHSDFYVMGFP-----GRLSNFGPADL 266  
 QY 294 -----EKGACGGG--GGHSEFQPOAVAVPAKR----- 319  
 Db 267 YSVQSRGPTPRPSNFENNAVKYGFYNTNTSSVPAAGSYPAFPNPFSTGTGVSTKPNKI 326  
 QY 320 -----KDLHMLVWSSASVPSERAHVHVFAGGADHADVLAKGAQA 360  
 Db 327 PKENQOQLQEKDASHDAKELHMFVWSSASVPSD-----VFGGAGDNV-----ATE 375  
 QY 361 YDEYGRDDY-----SSRTKNGSGGAD-----KGGPTLSKLGNSNTAQLY 399  
 Db 376 QSEQAKBEIRMVYSDQPRKSNARGGDDIGGLDSGEGEREIEKATAGLNKMGNSTAELE 435  
 QY 400 PKDDGEGRAAAVAMPASVMTLILIMVWKLTIRNNTYSSLLIGVWVSLVSYRWGIEMPA 459  
 Db 436 AAGDGGGNGGTMPPTSVMTLILIMVWKLTIRNNTYSSLLIGLWLVAYRWVHAMPK 495  
 QY 460 IIAARSISILSDAGLGMAMFSLGLFVALQPKIIACGNSVATFAMAVRFTITGPAIMAVAGIA 519  
 Db 496 ILQOSISILSDAGLGMAMFSLGLFVALQPKIIACGNSVATFAMAVRFTITGPAIMAVAGIA 555

QY 520 VGLRGVLLHIAIVQAAALPQGIIVFVFAKEYGVHPDILSTA--YG-----PIT 564  
 Db 556 IGLHGDLLRIAIVQAAALPQGIIVFVFAKEYGVHPDILSTA--YG-----PIT 607

## RESULT 11

AAI44270  
 ID AAY44270 standard; protein; 647 AA.

XX AC AAY44270;  
 XX  
 DT 28-FEB-2000 (first entry)  
 DE A. thaliana Ethylene Insensitive Root (EIR1) allele - EIR1-S97G.  
 XX  
 KW Root-specific auxin transport protein; Ethylene Insensitive Root;  
 KW gravitropism; auxin-based herbicide; plant root; transgenic plant;  
 KW auxin homeostasis; EIR1 allele; EIR1-S97G;  
 KW altered auxin-mediated response.  
 XX  
 OS Arabidopsis thaliana.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 97  
 FT /note= "Wild type Ser substituted by Gly"  
 XX  
 FN WO9963092-A1.  
 XX  
 PD 09-DEC-1999.  
 XX  
 PF 03-JUN-1999; 99WO-US012277.  
 XX  
 PF 03-JUN-1998; 98US-0087789P.  
 PR (WHED ) WHITEHEAD INST BIOMEDICAL RES.  
 XX  
 PI Luschning C, Gaxiola RA, Grisafi P, Fink GR;  
 XX WPI; 2000-086979/07.  
 DR  
 XX DNA encoding a root-specific auxin transport protein, used to develop  
 PT transgenic plants with increased resistance to herbicides.  
 XX  
 PS Claim 17; Page; 55pp; English.

CC The present sequence is an EIR1 allele which was obtained by replacing  
 the conserved Ser97 of A. thaliana EIR1 with Gly. EIR1 protein functions



QY 437 TYSSLGIVVMSLVSRWGIEMPAIARSISILSDAGLGMAFSLGLFMAIQPRIIACGNK 496  
Db 504 TYSSLFGLAWSLVSPKNIKMTIMSGSISILSDAGLGMAFSLGLFMAIQPRIIACGKS 563  
QY 497 LAATAMGRFVACAPVMAASIAVGLVILHIAVQAALPQGIIVFVFAKEYGVHPDIL 556  
Db 564 VAGFAMAVRFTGPAVIAATSAIGIRGDLHLHIAVQAALPQGIIVFVFAKEYGVHPDIL 623  
QY 557 STA--YG-----PIT 564  
Db 624 STAVIFGMLVALPVT 638

RESULT 13  
AAY44264  
ID AAY44264 standard; protein; 647 AA.

XX AC AAY44264;  
XX DT 28-FEB-2000 (first entry)  
XX DE A. thaliana Ethylene Insensitive Root (EIR1) protein.  
XX KW Root-specific auxin transport protein; EIR1; Ethylene Insensitive Root;  
XX KW gravitropism; auxin-based herbicide; plant root; transgenic plant;  
XX KW auxin homeostasis; plasma membrane; transmembrane domain.  
XX OS Arabidopsis thaliana.

XX FH Key Location/Qualifiers  
FT Peptide 1..60  
FT Domain /label= Signal\_peptide  
FT Domain 10..28  
FT Domain /label= Transmembrane\_domain  
FT Domain 41..60  
FT Domain /label= Transmembrane\_domain  
FT Domain 71..92  
FT Domain /label= Transmembrane\_domain  
FT Domain 100..121  
FT Domain /label= Transmembrane\_domain  
FT Domain 132..152  
FT Domain /label= Transmembrane\_domain  
FT Domain 503..520  
FT Domain /label= Transmembrane\_domain  
FT Domain 530..553  
FT Domain /label= Transmembrane\_domain  
FT Domain 568..589  
FT Domain /label= Transmembrane\_domain  
FT Domain 595..611  
FT Domain /label= Transmembrane\_domain  
FT Domain 622..642  
FT Domain /label= Transmembrane\_domain

XX PN WO9963092-A1.  
XX PD 09-DEC-1999.  
XX PF 03-JUN-1999; 99WO-US012277.  
XX PR 03-JUN-1998; 98US-0087789P.  
XX PA (WHED ) WHITEHEAD INST BIOMEDICAL RES.  
XX PI Luschnig C, Gaxiola RA, Grisafi P, Fink GR;  
XX DR WPI; 2000-086979/07.  
XX DR N-PSDB; AAZ29279, AAZ29278.  
XX PT DNA encoding a root-specific auxin transport protein, used to develop  
XX PT transgenic plants with increased resistance to herbicides.  
XX PS Claim 18; Fig 7; 55pp; English.

CC The present sequence is a EIR1 protein which functions as a root-specific  
CC auxin transport (efflux) carrier involved in gravitropism. The sequence  
CC was obtained from A. thaliana cDNA phage library. EIR1 consists of  
CC potential N-glycosylation sites and a N-terminal signal peptide which  
CC indicates localisation of the protein in the plasma membrane. The  
CC transmembrane domains are located in the highly conserved portions of the  
CC protein, N- and C-termini. The sequence is used for producing  
CC genetically engineered plants with greater resistance to auxin-based  
CC herbicides and auxin transport inhibitors in combination with a second  
CC herbicide. It can also be used to enhance transport of auxin in plant  
CC roots, produce transgenic plants which exhibit altered auxin homeostasis  
CC and mutant plants in which the roots are agravitropic and have reduced  
CC sensitivity to ethylene  
XX SQ Sequence 647 AA;

Query Match 52.7%; Score 1535; DB 3; Length 647;  
Best Local Similarity 51.4%; Pred. No. 5.6e-135;  
Matches 347; Conservative 64; Mismatches 116; Indels 148; Gaps 17;  
QY 1 MITALDLYHVLTAIVVLYVAMTAYGVSRWRIPTDQCSGINRFVALFAVLLSHPHFS 60  
Db 1 MITGKMDYDLAAWVLYVAMTAYGVSRWRIPTDQCSGINRFVALFAVLLSHPHFS 60  
QY 61 TNDPFAMNTRFLAADTLQKVAVLALL-----ALASRGLSSPRALGLDWSITLPSLSLTPNT 116  
Db 61 SNOPYANVYTHFLAADSLOKVILAAFLWQAFSRQG-----SLEWMTLISLTPNT 113  
QY 117 LVMGIPILLRNYGASSAGITLVQVIVVQLIITLMLFLFEYRAARALVLDQPDGAAS 176  
Db 114 LVMGIPILLRNYGDFS-GNLMVQIVVQLIITLMLFLFEYRAARALVLDQPDGAAS 171  
QY 177 IVSPRVDSVVSLL-ARGDVELEAEPDGVAGAGAVSRGSDAGRVVTVKSTSSSEAC 235  
Db 172 ITSFRVDSVLSLNGREPLQTDAEI-----GDDGKLHVVRSSAASSMISS 218  
QY 236 ---SHS---HSQTMQPRVNSLGVETIYSLQSRNPTPGSSFNHADFF----- 277  
Db 219 FNKSHGGGLNSMITPRASNLGVEIYVQSSREPTPRASSFNQTD FYAMFNASKAPSR 278  
QY 278 ---NIVGAAAKGGG----- 288  
Db 279 HGYTNSYGGAGAGPGDGVYLSQSSKGVTPRTSNFDEVMKTAAGRGGRSGSLYNN 338  
QY 289 -----GAAGDEKACGGGGGHSQPQAVAVPAKXKDLHMLVWSSSA 331  
Db 339 SVPSYPPNPMTFTGSGASGVKKESGGGSGG-----GVGVGQNKEMKMFVWSSSA 392  
QY 332 SPVSERAAVHVFAGGA-----DHADVLAKGAQAYDEY---GRDDYSRTKNG 376  
Db 393 SPVSEANAKNATRGSGSDVSTDPKVSIPPHDNLATKAMQNLINMSPGKGVEMDQDG 452  
QY 377 SGGADKGGPTLSKLGSTAQLYPKDDGGRAAAAMPPASVMTLILIMVWRKLRNPN 436  
Db 453 NNGGK--SPYMGKKGSD-----VEDGGFGRKQMPASPASVMTLILIMVWRKLRNPN 503  
QY 437 TYSSLGIVVMSLVSRWGIEMPAIARSISILSDAGLGMAFSLGLFMAIQPRIIACGNK 496  
Db 504 TYSSLFGLAWSLVSPKNIKMTIMSGSISILSDAGLGMAFSLGLFMAIQPRIIACGKS 563  
QY 497 LAATAMGRFVACAPVMAASIAVGLVILHIAVQAALPQGIIVFVFAKEYGVHPDIL 556  
Db 564 VAGFAMAVRFTGPAVIAATSAIGIRGDLHLHIAVQAALPQGIIVFVFAKEYGVHPDIL 623  
QY 557 STA--YG-----PIT 564  
Db 624 STAVIFGMLVALPVT 638

RESULT 14  
AAY44271  
ID AAY44271 standard; protein; 647 AA.  
XX

AC AAY44271;  
DT 26-FEB-2000 (first entry)  
DE A. thaliana Ethylene Insensitive Root (EIR1) allele - EIR1-S97A.  
KW Root-specific auxin transport protein; Ethylene insensitive Root;  
KW gravitropism; auxin-based herbicide; plant root; transgenic plant;  
KW auxin homeostasis; EIR1 allele; EIR1-S97A;  
KW altered auxin-mediated response.  
XX  
XX Arabidopsis thaliana.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT Misc-difference 97 /note= "Wild type Ser substituted by Ala"  
FT  
FN WO9963092-Al.  
XX  
XX 09-DEC-1999.  
XX  
XX 03-JUN-1999; 99WO-US012277.  
XX  
XX 03-JUN-1998; 98US-0087789P.  
XX  
XX (WHEED) WHITEHEAD INST BIOMEDICAL RES.  
XX  
XX Luschnig C, Gaxiola RA, Griseafi P, Pink GR;  
PI  
DR WPI; 2000-086979/07.  
XX  
XX  
XX DNA encoding a root-specific auxin transport protein, used to develop  
XX transgenic plants with increased resistance to herbicides.  
XX  
XX Claim 17; Page; 55pp; English.  
XX  
XX The present sequence is an EIR1 allele which was obtained by replacing  
XX the conserved Ser97 of A. thaliana EIR1 with Ala. EIR1 protein functions  
XX as a root-specific auxin transport (efflux) carrier and is involved in  
XX gravitropism. The EIR1 allele confers altered auxin-mediated responses in  
XX plants. The allele can be used to produce transgenic plants with high  
XX sensitivity to fluorinated indolic compounds, and mutant plants in which  
XX the roots are agravitropic and have reduced sensitivity to ethylene. EIR1  
XX the roots are agravitropic and have reduced sensitivity to ethylene. EIR1  
XX sequence is used for producing genetically engineered plants which have  
XX greater resistance to auxin-based herbicides and auxin transport  
XX inhibitors, and exhibit altered auxin homeostasis. Note: The present  
XX sequence is not shown in the specification but has been derived from EIR1  
XX protein sequence shown in figure-7  
XX  
XX Sequence 647 AA;  
Query Match 52.7%; Score 1535; DB 3; Length 647;  
Best Local Similarity 51.4%; Pred. No. 5.6e-135;  
Matches 347; Conservative 64; Mismatches 116; Indels 148; Gaps 17;  
QY 1 MITALDLHLVLTAVPLVYVMTLAYSVRWRIFTPDQCNGINRVALFAVPLLSFHFIS 60  
DB 1 MITGKMDYDLVLAAMVPLVYVMTLAYSVRWGIGFTPDQCNGINRVALFAVPLLSFHFIS 60  
QY 61 TNDPFAANLRLAADTLQKAVIALLL---ALASRGLSSPRALGLDWSITLFSLTPLNT 116  
DB 61 SNDPFAAMYHFLAADSLOKVVILALFLWQAFSRRG-----ALEWMITLFSLTPLNT 113  
QY 117 LVNGIPLLRGMTGASAGTLMVQVVLQCIHWITLMLFLFEVRAARALVLPQPDGAAS 176  
DB 114 LVNGIPLLRAMTGDFS-GNLAMVQIVVLQSIHWITLMLFLFEFRGAKLLISQFPE-TAGS 171  
QY 177 IYSFRVDSVSVSI-ARGDVELEAE PDGVAGAGAVSSRGDAGRVTIVRKTSRSEAC 235  
DB 172 IYSFRVDSVLSLNGREPLQDAEI-----GDDGKLHVVRSSAASSMISS 218  
QY 236 ---SHS---HSQTMQPRVSNLSGVLEIYSLSQSRNTPRGSSFNHADFF----- 277

DB 219 FNKSHGGGLNSMITPRASNLTVGVIEYSVQSSREPTPRASSFNQDFYAMFNASKAPSPR 278  
QY 278 ----NIVGAAAKGG----- 288  
DB 279 HGYTNSYGGAGAGPGGDVYLSQSGVTPRTSNPDEEVWTKAKKAGRGRSGSLYNNN 338  
QY 289 -----GAAGDEEKACGCGGGGHSPOQAVAVPAKRDLEMLWSSSA 331  
DB 339 SVPSYPPNPMPFTGSTGASGVKKESGGSGG-----GVGVGQNKEMMFVWSSSA 392  
QY 332 SPVSEAAVHVFGAGGA-----DHADVLAKGAQAYDEY---GRDDYSSTRTKG 376  
DB 393 SPVSEAAKAVNTRGSSTDVSTDPKVSIPPHDNLATKAMQNLINNSPGRKHVEMDQDG 452  
QY 377 SGADKGGGPTLSKLGNSSTAQLYPKDDGEGRAAAVAMPASPVMTRLILIMVWRKIRNPN 436  
DB 453 NNGGK--SPYMGKGGSD-----VEDGGPGPRKQMPASPVMTRLILIMVWRKIRNPN 503  
QY 437 TYSSLLGVVWSLVSRWGIEMPAITARSISILSDAGLGNAMESLGLFWALQPRILACGNK 496  
DB 504 TYSSLFGLAWSLVSPKWNIMPTIMSGSISILSDAGLGNAMESLGLFWALQPRILACGKS 563  
QY 497 LAALAMGVRFVAGPVAAMAAASIAVGLRGLVLLHIAIVQALPQGIYVFFVFAKEYGVHPDIL 556  
DB 564 VAGFAMAVRFLTGPAVIAATSIAGIRGDLHLIAIVQALPQGIYVFFVFAKEYNVHPDIL 623  
QY 557 STA--YG-----PIT 564  
DB 624 STAVIEGMLVALPVT 638  
RESULT 15  
AAB26949  
ID AAB26949 standard; protein; 647 AA.  
XX  
XX AAB26949;  
XX  
XX 02-FEB-2001 (first entry)  
XX Auxin transport protein EIR1.  
XX  
XX Auxin transport protein; corn; root development; gene mapping;  
XX Plant breeding; herbicide.  
XX Arabidopsis thaliana.  
XX WO200068389-A2.  
XX 16-NOV-2000.  
XX  
XX 03-MAY-2000; 2000WO-US012061.  
XX  
XX 07-MAY-1999; 99US-0133040P.  
XX  
XX (DUPO) DU PONT DE NEMOURS & CO E I.  
XX (PION-) PIONEER HI-BRED INT INC.  
XX Orozco EM, Weng Z, Bruce WB, Cahoon RE, Tao Y;  
XX WPI; 2000-687647/67.  
XX  
XX New nucleic acid sequences encoding new auxin transport proteins, useful  
XX for modulating root growth of plants and to screen for herbicides.  
XX  
XX Disclosure; Fig 1; 94pp; English.  
XX  
XX Auxins are plant hormones that influence plant behaviour and development  
XX e.g. vascular tissue differentiation, apical development, tropic  
XX responses and organ (e.g. flower, leaf) development. The present  
XX invention relates to auxin transport proteins and coding sequences. The  
XX present sequence is one such auxin transport protein. The DNA sequence  
XX encoding this protein may be used to modulate root development, e.g. to

CC produce a more robust root system, alter root angle or redirect root  
CC growth; and may be useful for gene mapping (e.g. for plant breeding) and  
CC to identify loss of function mutants. This protein may be useful for  
CC raising specific antibodies, for the detection of auxin transport  
CC proteins and to design and/or identify specific inhibitors of auxin  
CC transport proteins, potentially useful as herbicides  
XX  
SQ Sequence 647 AA;

Query Match	52.7%; Score 1535; DB 3; Length 647;
Best Local Similarity	51.4%; Pred. No. 5.6e-135;
Matches 347; Conservative 64; Mismatches 116; Indels 148; Gaps 17;	
QY	1 MITALDLHYLTAVVPLVYAVWTLAYGSRVWEIFTPDOCSGINRVALFAVPLLSFHFIS 60
DB	1 MITGDMYDVLAAVPLVYAVWTLAYGSRVWEIFTPDOCSGINRVALFAVPLLSFHFIS 60
QY	61 TNDPFAMNLFADTLQKAVLALL----ALASRGLSPRALGLDMSITLFSLTLPNT 116
DB	61 SNDPYAMNVHFLAADSLOKXVILAAFLWQAFSRG-----SLEWMITLFSLTLPNT 113
QY	117 LVNGIPLLRGMYGASSAGTLVQVVLQCIWYTLMLFLFEYRARALVLDQPDGAAS 176
DB	114 LVNGIPLLRGMYGASSAGTLVQVVLQCIWYTLMLFLFEYRARALVLDQPDGAAS 176
QY	177 IVSFRVDSVWSL-ARGDVELEAEPDGVAGAVSSRGDAGRVRVTVRKSTSRSEAAC 235
DB	172 ITSFRVDSVWSL-ARGDVELEAEPDGVAGAVSSRGDAGRVRVTVRKSTSRSEAAC 235
QY	236 ----SHS----HSQTMOPRVSNLGVEIYLSQSRNTPRGSSFNHADFF----- 277
DB	219 FNKSHGGGLNSMITPRASNLTGVEIYVSQSRNTPRGSSFNHADFF----- 277
QY	278 ----NIVGAAAKGG----- 288
DB	279 HGVTNSYGAGAGPGGDVYVSLOSSKGVTPRTSNPDEEVMTAKKAGRGSRMSGELYN 338
QY	289 -----GAAGDEEKACGGGGHSPQPAVAPAKKDLHMLVWSSA 331
DB	339 SVPSYPPNPMTFTGSTSGASGVKKESGGSGG-----GVGVGGQNKEMMFVWSSA 392
QY	332 SPVSEBAVHVFGAG-----DHADVLAKGAQAYDEY---GRDDYSSTKNG 376
DB	393 SPVSEBAVHVFGAG-----DHADVLAKGAQAYDEY---GRDDYSSTKNG 376
QY	377 SGADKGGPTLSKLSNSTAQLYPKDDGGRGAAVAMPPEASVMTLRLIIMVWRKLIRPN 436
DB	453 NNGGK--SPYMGKGGD-----VEDGGPGPKQMPPEASVMTLRLIIMVWRKLIRPN 503
QY	437 TYSSLIQVWVSLVSYRWGTEMPAIIARSISILSDAGLGNMFSLGLFMALQPIIACGNK 496
DB	504 TYSSLIQVWVSLVSYRWGTEMPAIIARSISILSDAGLGNMFSLGLFMALQPIIACGNK 496
QY	497 LAAIAMGVRFVAGPVMYMAASIAVGLRGLLHIAIVQAAALPQGIVPFVFAKEYGVHPDIL 556
DB	564 VAGFANAVRFLTGPAVIATSTAIIRGDLHIAIVQAAALPQGIVPFVFAKEYGVHPDIL 623
QY	557 STA--YG-----PIT 564
DB	624 STAVIFGLVALPVT 638

Search completed: March 3, 2004, 09:07:00  
Job time : 80 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: March 3, 2004, 09:09:24 ; Search time 608 Seconds  
(without alignments)  
4003.647 Million cell updates/sec

Title: US-10-030-884-14  
Perfect score: 2910  
Sequence: 1 MITALDIYHVTAVPPIYVA.....DILSTAYGPTSHGFITCHS 573

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Dgapop 6.0 , Delcxt 7.0

Searched: 337863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame\_plus\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_Spool/US10030884/runat\_24022004\_103922\_1940/app\_query.fasta\_1.711  
-DB=N Geneseq\_29Jan04 -QMT=fascap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=DIOSUM62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10030884 @CGN 1.1 470 @runat\_24022004\_103922\_1940 -NCPU=6 -ICPU=3  
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

- Database : N\_Geneseq\_29Jan04:\*
- 1: Geneseqn1980s:\*
  - 2: Geneseqn1990s:\*
  - 3: Geneseqn2000s:\*
  - 4: Geneseqn2001as:\*
  - 5: Geneseqn2001bs:\*
  - 6: Geneseqn2002as:\*
  - 7: Geneseqn2003as:\*
  - 8: Geneseqn2003bs:\*
  - 9: Geneseqn2003cs:\*
  - 10: Geneseqn2004as:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2910	100.0	2769	AAA94721	AAA94721 Corn auxi
2	2129.5	73.2	1926	ADA70391	Ada70391 Rice gene
3	2020.5	69.4	1686	ADA69470	Ada69470 Rice gene
4	1665	57.2	2374	AAZ29280	AAZ29280 Rice EIR1
5	1647.5	56.6	1827	ADA70668	Ada70668 Rice gene
6	1643	56.5	2549	AAA94729	AAA94729 Soybean a
7	1641.5	56.4	2292	AAV16361	AAV16361 cDNA sequ
8	1596	54.8	1845	ADA69396	Ada69396 Rice gene

9	1592	54.7	1884	7	ADA70233	Ada70233 Rice gene
10	1566.5	53.8	2293	3	AAA94733	AAA94733 Wheat aux
11	1560.5	53.6	1923	3	AAZ44722	AAZ44722 Arabidops
12	1560.5	53.6	1923	6	ABZ13042	ABZ13042 Arabidops
13	1542.5	53.0	1851	3	AAZ43229	AAZ43229 Arabidops
14	1535	52.7	2204	3	AAZ29279	AAZ29279 A. thalia
15	1535	52.7	2232	3	AAZ57348	AAZ57348 Gravitrop
16	1527	52.5	2324	6	AAA94731	AAA94731 Soybean a
17	1526	52.4	1860	6	ABZ13043	ABZ13043 Arabidops
18	1526	52.4	1860	7	ADA68467	Ada68467 Arabidops
19	1526	52.4	2027	3	AAZ44892	AAZ44892 Arabidops
20	1513	52.0	2162	3	AAA94724	AAA94724 Rice auxi
21	1408.5	48.4	1725	3	AAZ46518	AAZ46518 Arabidops
22	1406.5	48.3	1983	6	ABZ13500	ABZ13500 Arabidops
23	1338	46.0	2117	2	AAV16362	AAV16362 cDNA sequ
24	1318.5	45.3	2061	3	AAZ46177	AAZ46177 Arabidops
25	1299.5	44.7	2101	3	AAA94727	AAA94727 Soybean a
26	1228	42.2	7072	3	AAZ29278	AAZ29278 A. thalia
27	1218.5	41.9	3980	3	AAZ57349	AAZ57349 Gravitrop
28	1183	40.7	1713	3	AAZ48532	AAZ48532 Arabidops
29	1004	34.5	1426	3	AAA94719	AAA94719 Corn auxi
30	965	33.2	1618	3	AAA94725	AAA94725 Soybean a
31	662.5	22.8	1088	3	AAA94716	AAA94716 Corn auxi
32	630	21.6	855	3	AAA94737	AAA94737 Corn auxi
33	454	15.6	504	3	AAA94720	AAA94720 Corn auxi
34	385	13.2	425	3	AAA94736	AAA94736 Wheat aux
35	383.5	13.2	415	3	AAA94735	AAA94735 Wheat aux
36	356	12.2	473	3	AAA94732	AAA94732 Wheat aux
37	334	11.5	330	3	AAA94723	AAA94723 Rice auxi
38	323	11.1	531	3	AAA94726	AAA94726 Soybean a
39	322	11.1	419	3	AAA94730	AAA94730 Soybean a
40	320.5	11.0	543	3	AAA94722	AAA94722 Rice auxi
41	316	10.9	624	3	AAA94718	AAA94718 Corn auxi
42	303.5	10.4	629	3	AAA94715	AAA94715 Corn auxi
43	273	9.4	525	3	AAA94728	AAA94728 Soybean a
44	259.5	8.9	447	3	AAA94734	AAA94734 Wheat aux
45	257	8.8	620	7	ABZ73099	ABZ73099 Rice leaf

ALIGNMENTS

RESULT 1  
AAA94721  
ID AAA94721 standard; DNA; 2769 BP.  
XX AC AAA94721;  
XX DT 02-FEB-2001 (first entry)  
XX DE Corn auxin transport protein clone p0119.cmtnl24r DNA sequence.  
XX KW Auxin transport protein; corn; root development; gene mapping;  
XX OS plant breeding; herbicide; ss.  
XX OS Zea mays.  
XX PN WO200068389-A2.  
XX PD 16-NOV-2000.  
XX PF 03-MAY-2000; 2000WO-US012061.  
XX PR 07-MAY-1999; 99US-0133040P.  
XX PA (DUPO) DU PONT DE NEMOURS & CO E I.  
(PTON-) PIONEER HI-BRED INT INC.  
XX PI Orozco EM, Weng Z, Bruce WB, Cahoon RE, Tao Y;  
XX DR WPI; 2000-687647/67.  
XX DR P-PSDB; AAB26934.  
XX PT New nucleic acid sequences encoding new auxin transport proteins, useful

PT for modulating root growth of plants and to screen for herbicides.  
XX Claim 2; Page 60-61; 9app; English.  
XX Auxins are plant hormones that influence plant behaviour and development  
CC e.g. vascular tissue differentiation, apical development, tropic  
PS responses and organ (e.g. flower, leaf) development. The present  
XX invention relates to corn auxin transport protein coding sequences and  
CC proteins. The present sequence is one such coding sequence. This sequence  
CC may be used to modulate root development, e.g. to produce a more robust  
CC root system, alter root angle or redirect root growth. Also, the present  
CC sequence may be useful for gene mapping (e.g. for plant breeding) and to  
CC identify loss of function mutants. The protein encoded by the present  
CC sequence may be useful for raising specific antibodies, for the detection  
CC of auxin transport proteins and to design and/or identify specific  
XX inhibitors of auxin transport proteins, potentially useful as herbicides  
SQ Sequence 2769 BP; 527 A; 917 C; 757 G; 568 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 1.41e-194 Length: 2769  
Score: 2910.00 Matches: 573  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

US-10-030-884-14 (1-573) x AAA94721 (1-2769)

QY 1 MetIleThrAlaLeuAspLeuTyHisValLeuThrAlaValValProLeuTyValAla 20  
DB 171 ATGATCACCGGGTGGACCTTACACAGTGTGACGGCGGTGGTGGCGTGTACGTGGCC 230

QY 21 MetThrLeuAlaTyGlySerValArgTTPArgIlePheThrProAspGlnCysSer 40  
DB 231 ATGACGTGGGTACGGTCCGTCGGTGGTGGGCGATCTTACGCGGACAGTGTCTCC 290

QY 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
DB 291 GGGATCAACCGCTTTCGGTGGCTTTCGGCGCTTCGGCGCTTCTCTTCTTCTTCTTCT 350

QY 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80  
DB 351 ACCACGACCCCTTCGCATGAACTCGGCTTCTGTCGCGCCGACACGCTGCAGAGGTG 410

QY 81 AlaValLeuAlaLeuLeuAlaLeuAlaSerArgGlyLeuSerSerProArgAlaLeuGly 100  
DB 411 GCGGTCTCTGGCGTGTGCGGCTTCCGCGCTTCCGCGGCTTCTTCTTCCGCGCGCTCGG 470

QY 101 LeuAspTTPSerIleThrLeuPheSerLeuSerThrLeuProAsnThrLeuValMetGly 120  
DB 471 CTGACTGGAGCATCAGCTCTTCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 530

QY 121 IleProLeuLeuArgGlyMetTyGlyAlaSerSerAlaGlyThrLeuMetValGlnVal 140  
DB 531 ATCCCGCTGTCTGGAGGCGATGACGGCGCTGCTGCGCGCGCTTCTTCTTCCGCGCGCT 590

QY 141 ValValLeuGlnCysIleIleThrTyThrLeuMetLeuPheLeuPheGluTyArgAla 160  
DB 591 GTGCTCTTCCAGTGCATCATCTGTATACGCTCATGCTTCTTCTTCTTCTTCTTCTTCTT 650

QY 161 AlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSerIleValSerPhe 180  
DB 651 GCGCGCGGCTCGTCTTCTGACAGTTCCTCCGACGCGCGCGCTTCTTCTTCTTCTTCTTCT 710

QY 181 ArgValAspSerAspValValSerLeuAlaArgGlyAspValGluLeuGluAlaGluPro 200  
DB 711 GCGGTGACTCCGACGTCTCTGCTTCCGCGGCGGAGCTCGAGCTTCGAGCGCGCGGCTCC 770

QY 201 AspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGlyAspAlaGlyArgValArg 220  
DB 771 GACGCGGTGCGCGCGCGCGCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 830

QY 221 ValThrValArgLysSerThrSerSerArgSerGluAlaAlaCysSerHisSer 240  
DB 831 GTCACTGTGGCAAGTCCACAGCTCGCTCGAGGCGCGTGCAGCTCGCACTCC 890

QY 241 GlnThrMetGlnProArgValSerAsnLeuSerGlyValGluIleTySerLeuGlnSer 260  
DB 891 CAGACCATGTGAGCCCGCTGTGTCTCAACTCTCGGCGTGGAGATCTACTCGTGCAGTCG 950

QY 261 SerArgAsnProThrProArgGlySerSerPheAsnHisAlaAspPhePheIleVal 280  
DB 951 TCGGCGCAACCCACCCCGCGCGGTTCAGCTTCAACAGCCGACTTCTTCAACATCGTC 1010

QY 281 GlyAlaAlaLysGlyGlyGlyGlyAlaAlaGlyAspGluGluLysGlyAlaCysGly 300  
DB 1011 GCGCGCGCGCGCGAGGAGGCGCGGCGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1070

QY 301 GlyGlyGlyGlyGlyHisSerProGlnProGlnAlaValAlaValProAlaLysArgLys 320  
DB 1071 GCGCGCGCGGAGGAGACTCGCGCGAGCGCGAGCGCTGCGCGTGGCGCAAGAGAGAG 1130

QY 321 AspLeuHisMetLeuValTTPSerSerSerAlaSerProValSerGluArgAlaVal 340  
DB 1131 GACCTGCACATGCTGTGTGGAGCTCCAGCGCTTCGCGCGTGTTCGAGGCGCGCGCGTG 1190

QY 341 HisValPheGlyAlaGlyGlyAlaAspHisAlaAspValLeuAlaLysGlyAlaGlnAla 360  
DB 1191 CACGTCTTCGCGCGCGCGCGCTGACCATGCGAGCTCTCGCGCAAGAGCGCGCGCGCC 1250

QY 361 TyrAspGluTyGlyArgAspAspTySerSerArgThrIleAsnGlySerGlyAla 380  
DB 1251 TACGACGAGTACGCGCGCGAGCTACTACGAGCAGGAGCAAGAGAGCGCGCGCGCG 1310

QY 381 AspLysGlyGlyProThrLeuSerLysLeuGlySerAsnSerThrAlaGlnLeuTyPro 400  
DB 1311 GACAAAGCGCGCGCGCGCTGTGAGAGCTGGGGTCCAACTCGACGCGCGAGCTGTACCC 1370

QY 401 LysAspAspGlyGluGlyArgAlaAlaAlaValAlaMetProProAlaSerValMetThr 420  
DB 1371 AAGGACACGCGCGAGGCGCGCGCGGTGGCGATCCGCGCGCGAGCTGTATGACG 1430

QY 421 ArgLeuIleLeuIleMetValTTPArgLysLeuLeuLeuArgAsnProAsnThrTySerSer 440  
DB 1431 CGGCTCATCTCTCATCATGTGTGGAGAGAGCTGATCCGGAAACCCCAACACCTTACTCCAGC 1490

QY 441 LeuIleGlyValValTTPSerLeuValSerTyArgTTPGlyIleGluMetProAlaIle 460  
DB 1491 CTCATCGCGCTCGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 1550

QY 461 IleAlaArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeu 480  
DB 1551 ATGCGCGCGTTCGATTCGATCTTGTGAGCGCGGCTCTCGGGATGGCCATGTTTACGCTA 1610

QY 481 GlyLeuPheMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAlaIle 500  
DB 1611 GCGCTGTTCATGCGCTGCGCGCGAGGATCATCGCGTGGCGGAAACAAAGCTGCGCGCATC 1670

QY 501 AlaMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaAlaSerIleAlaVal 520  
DB 1671 GCGATGGCGCTCGGTTCGTTCGAGGCGCGCGGCTCATGCGCGCGCTTCCATCGCGCTC 1730

QY 521 GlyLeuArgGlyValLeuLeuHisIleAlaIleValGlnAlaAlaLeuProGlnIle 540  
DB 1731 GGTCTGCGCGCGCTCTCTCTCCCATCGCATCGCTGTCCAGCTGTCTGTCTGTCTGTCTGT 1790

QY 541 ValProPheValPheAlaLysGluTyGlyValHisProAspIleLeuSerThrAlaTy 560  
DB 1791 GTGCGCTTCGTTCGCGCAAGGAGTACGGCGTTCATCCCGACATCTCTGAGCAGCGAT 1850

QY 561 GlyProIleThrSerHisGlyPheIleThrCysHisSer 573  
DB 1851 GGTCCAATAACATCGCATGTTTTCATCTTCTTCCATAGT 1889

RESULT 2





XX The present invention relates to a method (M1) for identifying genes  
CC involved in plant resistance or response to pathogenic infection. M1  
CC comprises identifying a gene whose expression is significantly altered in  
CC the incompatible interaction of plant gene expression relative to  
CC expression of the gene in an uninfected plant, in a mutant plant that  
CC does not express a gene associated with response to pathogenic infection,  
CC or in a corresponding incompatible or compatible interaction. (M1) is  
CC useful for conferring resistance to resistance or tolerance to a plant to  
CC bacterial, fungal or viral infection. The present sequence was used to  
CC illustrate the invention.

XX  
SQ Sequence 1686 BP; 286 A; 553 C; 490 G; 356 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.:	1.81e-132	Length:	1686
Score:	2020.50	Matches:	442
Percent Similarity:	75.69%	Conservative:	22
Best Local Similarity:	72.10%	Mismatches:	39
Query Match:	69.43%	Indels:	111
DB:	7	Gaps:	16

US-10-030-884-14 (1-573) x ADA69470 (1-1686)

QY	1	MettLeThrAlaLeuAaspLeuTyrHisValLeuThrAlaValProLeuTyrValAla	20
DB	1	ATGATCAAGGTGGTGGACCTGTATACAGCTCTCAGCGCGTGTGCGGTGTGATCGTGGCG	60
QY	21	MetThrLeuAlaTyrGlySerValAagTTrpTArgIlePheThrProAepGlnCysSer	40
DB	61	ATGACGCTGGCGGTACGGCTCCGTCCGTGGTGGCGCATCTTCTCCCGGACGAGTCTCC	120
QY	41	GlyLeuAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer	60
DB	121	GGCATCAACCGTTCGTCCGGCTCTTCGCGCTCCGCTCTCTCTCCATTCATCTCTCC	180
QY	61	ThrAsnAapProPheAlaMetAsnLeuArgPheLeuAlaAaAaspThrLeuGlnLysVal	80
DB	181	ACCAACAACCCCTTCGCCATGAACTCCGCTTCTCGCCGCCACGACGCTCCAGAGCTC	240
QY	81	AlaValLeuAlaLeuLeuAla-----LeuAlaSerArgGlyLeuSerSerPro	96
DB	241	ATCGTCTCGCCCTCTTCGGCTCTGTGGCGCTCTCTCCGCGCGCGG-----	288
QY	97	ArgAlaLeuGlyLeuAaspTrpSerIleThrLeuPheSerLeuSer-----ThrLeuProAsn	116
DB	289	-----TCCCTCGACTGGCTCATCACCTCTTCTCCCTTCCCAACTTCCCAAMAC	339
QY	116	rLeuValMetGlyIleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLe	136
DB	340	CCTCGTCATGGCATCCGCTGCTCAAGGGAT-----	372
QY	136	uMetValGlnValValValLeuGlnCysIleIleTrpTyrThrLeuMetLeuPheLeuPh	156
DB	373	-----TGCATCATCTGGTACACGCTCATGCTGTGTCTCTCT	407
QY	156	eClyThrArgAlaAlaArgAlaLeuValLeuAaspGlnPheProAaspGlyAlaAlaAla	176
DB	408	CGAGTACCGTGGCGCGCGCTTCTGTGATGGAGCAGTCTCCCGAC-----ACCGCGCCTC	464
QY	176	rIleValSerPheArgValAaspSerAaspValValSerLeuAlaAargGlyAasp-----	193
DB	465	CATGCTCTCTTCGGGTGGATTCCGACGTCGTCGCTCGCGCGCGCGCGTGTGTGGAGC	524
QY	194	-ValGluLeuGluAlaGluProAaspGlyValAlaGlyAlaValSerSerArgG1	213
DB	525	CGCGAGCTGACGCGGAG-----SerHisSerGlnThrMetGlnProArgValSerAsnLe	250
QY	213	yGlyAaspAlaGlyArgValArgValThrValArgLysSerThrSerSerArgSerGluAl	233
DB	552	CGGGACCATGGCAAGATCGGAGTACCGTGGCAAGTCGACGAGCTCGGCTCCGAGGC	611
QY	233	aAlaCysSerHis-----SerHisSerGlnThrMetGlnProArgValSerAsnLe	250

Db 612 GCGCTGCTCGCAGCGGACGAGTCCACACTCGAGTCCATGCGCGCGCTCCAACT 671  
 QY 250 userGlyValGluIleuTySerLeuGlnSerArgAsnProThrProArgGlySerSe 270  
 Db 672 CTCGGGCTCGAGATTACTCGCTCGAGTCTCGCGAACCCGACCGCGCGGCTCCAG 731  
 QY 270 rPheAsnHisAlaAspPhePheAsnIleValGlyAlaAlaAlaGlyGlyGlyAl 290  
 Db 732 CTTCAACACACCGCGAGTCTTCAACATCGTC-----GGCAACGGCAA 773  
 QY 290 aAlaGlyAspGluGluGlyAlaCysGlyGlyGlyGlyGlyGlyHisSerProGlnPr 310  
 Db 774 GCACGGGACAGAGAGAGGCGGC-----GCCGGTGGCGGGGCGCACCTCGCGCAGCC 827  
 QY 310 cGlnAlaValAlaValProAlaLysArgLysAspLeuHisMetLeuValTrpSerSerSe 330  
 Db 828 G-----GTGGTGGGGAAGAGAGGACCTGCACATGTCGTGGAGCTCAAG 875  
 QY 330 rAlaSerProValSerGluArgAla-----AlaValHisValPh 343  
 Db 876 CGCTCGCGGTGTCGAGCGCGCGCGCGCGCGCTCGCGCGCGTGCACGTCCT 935  
 QY 343 eGlyAlaGlyGlyAlaAspHisAlaAspValLeuAlaLysGlyAlaGlnAlaTyAspGl 363  
 Db 936 CGGCGGTGTCGCGCGACCGACGCGAC-----GCCAAGGTGCTCAGGCTATGATGA 989  
 QY 363 uTyzGlyArgAspAspTySerSerArgThrLysAsnGlySerGlyGlyAlaAspLysGl 383  
 Db 990 G-----TACAGCTTCGGGAACAGAAAT-----GAGAAGGA 1019  
 QY 383 yGlyProThrLeuSerLysLeuGlySerAsnSerThrAlaGlnLeuTyProLysAspAs 403  
 Db 1020 CGGCGCGACATGTCGAAGCTGGGGTCCACTCGACGGCGGAGCTCCGCGCAAGGACGA 1079  
 QY 403 pGlyGluGlyArgAlaAlaAlaValAlaMetProProAlaSerValMetThrArgLeuIl 423  
 Db 1080 CGGCGAGGGAGGGCGCA-----GCGATCGCGCGCGAGCGTGATGACGAGGCTCAT 1133  
 QY 423 eLeuIleMetValTrpArgLysLeuIleArgAsnProAsnThrTySerSerLeuIleGl 443  
 Db 1134 ACTGATCATGCTCTGGAGAGAGTGCATCAGAACCCAAACACTTACTCCAGCCTCTCGG 1193  
 QY 443 yValValTrpSerLeuValSerTyArgTrpGlyIleGluMetProAlaIleIleAlaAr 463  
 Db 1194 TGTCATCTGGTCCCTCTCTCTACAGTGGGGATTCAGATGCGGCTATCATCGCCCG 1253  
 QY 463 gSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeuGlyLeuPh 483  
 Db 1254 GTCGATTTTCGATCTCTTCAGATGCGGGCTCGGAATGCGCATGTTCCAGCCTTAGGATTGT 1313  
 QY 483 eMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAlaIleAlaMetGl 503  
 Db 1314 CATGCATTGCAGCCAGCATTCCTCTGGGACCTCCCTTGTGATGCAATGGC 1373  
 QY 503 yValArgPheValAlaGlyProAlaValMetAlaAlaSerIleAlaValGlyLeuAr 523  
 Db 1374 CGTCAGGTTCTCGTGTGCTCGCTCATGCTCGCGCTCCATCGCGCTCGGACTTCG 1433  
 QY 523 gGlyValLeuLeuHisIleAlaIleVal----- 532  
 Db 1434 CGGGTGTCTTCGACATTGCATTGTTTCAGTAAGTGCAGTTCAGGCTTCGACCA 1493  
 QY 533 ----- 533  
 Db 1494 AAATGCATCTTAACCTGCAAGAAATCTATCGTAACCAACATTTGTCATTGAATCTCA 1553  
 QY 533 nAlaAlaLeuProGlnGlyIleValProPheValPheAlaLysGlnTyArgValHisPr 553  
 Db 1554 GGCGGCTCTTCCTCAAGGAATCGTCCCTTCGTGTTTGCAGGAGTACATATGTTTCATCC 1613  
 QY 553 oAspIleLeuSerThrAlaTyArgGlyProIleThrSer 565  
 Db 565

Db 1614 TAACATTCTGAGCACAGCGTATGATCTCTTTTACCACC 1650  
 RESULT 4  
 AA229280  
 ID AAZ29280 standard; cDNA; 2374 BP.  
 XX  
 AC AAZ29280;  
 XX  
 DT 28-FEB-2000 (first entry)  
 XX  
 DE Rice EIR1 Homologue (REH1) cDNA.  
 XX  
 KW Root-specific auxin transport protein; EIR1; Ethylene Insensitive Root;  
 KW gravitropism; auxin-based herbicide; plant root; transgenic plant;  
 KW auxin homeostasis; Rice EIR1 Homologue; REH1; ds.  
 XX  
 OS Oryza sativa.  
 XX  
 PH Key Location/Qualifiers  
 FT CDS 158..1945  
 FT /tag= a  
 FT /product= "REH1 protein"  
 FT sig\_peptide 158..337  
 FT /tag= b  
 XX  
 PN WO9963092-Al.  
 XX  
 PD 09-DEC-1999.  
 XX  
 PF 03-JUN-1999; 99WO-US012277.  
 XX  
 PR 03-JUN-1998; 98US-0087789P.  
 XX  
 PA (WHED) WHITEHEAD INST BIOMEDICAL RES.  
 XX  
 PI Luschnig C, Gaxiola RA, Grisafi P, Fink GR;  
 XX  
 DR WPI; 2000-086979/07.  
 DR P-PSDB; AAV44265.  
 XX  
 PT DNA encoding a root-specific auxin transport protein, used to develop  
 PT transgenic plants with increased resistance to herbicides.  
 XX  
 PS Claim 1; Fig 8; 55pp; English.  
 XX  
 CC The present sequence is a cDNA encoding REH1, a rice homologue of EIR1  
 CC which functions as a root-specific auxin transport (efflux) carrier  
 CC protein involved in gravitropism. The sequence is obtained from a rice  
 CC EST derived from root-specific cDNA. The sequence is used for producing  
 CC genetically engineered plants with greater resistance to auxin-based  
 CC herbicides and auxin transport inhibitors in combination with a second  
 CC herbicide. It can also be used to enhance transport of auxin in plant  
 CC roots, produce transgenic plants which exhibit altered auxin homeostasis  
 CC and mutant plants in which the roots are agravitropic and have reduced  
 CC sensitivity to ethylene  
 XX  
 SQ Sequence 2374 BP; 471 A; 736 C; 733 G; 434 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 2,26e-107 Length: 2374  
 Score: 1665.00 Matches: 365  
 Percent Similarity: 67.58% Conservative: 54  
 Best Local Similarity: 58.87% Mismatches: 111  
 Query Match: 57.22% Indels: 90  
 DB: 3 Gaps: 17  
 US-10-030-884-14 (1-573) x AAZ29280 (1-2374)  
 QY 1 MetIleThrAlaLeuAspLeuTyHisValLeuThrAlaValProLeuTyValAla 20  
 Db 158 ATGATTACGCGCGCGACTTCTTACACGATGACGCGCGATGTCGCTGTGACGTGGCG 217  
 QY 21 MetThrLeuAlaTyArgSerValArgTrpArgIlePheThrProAspGlnCysSer 40

Db 218 ATGATACGGCGTACGGTGGTGAAGTGGTGGCGCATCTTCACGGCGGACGATGCTCC 277  
 Qy 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
 Db 278 GGGATCAACCGCTTCGTGGCGCTCTTCGCGGTGCGCTGCTGCTTCACATTCATCTCC 337  
 Qy 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80  
 Db 338 ACCAACACCCGTACAGATGAACCTTCGGTTCATCGCGCGGACACGCTGCAGAGCTG 397  
 Qy 81 AlaValLeuAlaLeuAla-----LeuAlaSerArgGlyLeuSerSerPro 96  
 Db 398 ATGGTGTGGCGCATGCTACGGCGGTGGAGCCACCTTCAGCGCGCGGGG----- 445  
 Qy 97 ArgAlaLeuGlyLeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThr 116  
 Db 446 -----AGCCTCGAGTGAGCAATACGCTCTTCCTCTCCACGCTGCCCAACACG 496  
 Qy 117 LeuValMetGlyIleProLeuLeuArgGlyMetTyArgIleAlaSerSerAlaGlyThrLeu 136  
 Db 497 CTCGTATGGGATCCCTTTGTCTCAAGGCGCATGTACGGGAGTTCTCC---GGCAGCTC 553  
 Qy 137 MetValGlnValValValLeuGlnCysIleIleTrpThrLeuMetLeuPheLeuPhe 156  
 Db 554 ATGGTGCAGATCGTGTGTCAGTGCAATCATCTGGTACACGCTCATGCTCTTCATGTT 613  
 Qy 157 GluTyArgAlaAlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSer 176  
 Db 614 GAGTACCGCGCGCGCGATGCTCATCCAGGAGCAGTTCCTCCGAC---ACGCGCGCAAC 670  
 Qy 177 IleValSerPheArgValAspSerAspValValSerLeu-----AlaArgGlyAspVal 194  
 Db 671 ATCGCCTCCATCGTCTGACCGCGGACGTCGTGTCGTCGACGCGGAGGAGGCCCATC 730  
 Qy 195 GluLeuGluAlaGluProAspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGly 214  
 Db 731 GAGACCGAGCGAGGTGAAG----- 751  
 Qy 215 AspAlaGlyArgValArgValThrValArgLysSerThrSerSerArgSerGluAlaAla 234  
 Db 752 GAGGACCGGAGGATACATCACTCCGTCGCGCGCTCCACACGCTCTCGTTCGACATCTAC 811  
 Qy 235 CysSerHisSerHisSer---GlnThrMetGlnProArgValSerAsnLeuSerGlyVal 253  
 Db 812 TCCCGCGCTCCATGGGCTTCCACACACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 871  
 Qy 254 GluIleTySerLeuGlnSerArgAsnProThrProArgGlySerSerPheAsnHis 273  
 Db 872 GAGATCTACTCGTCTGAGTCTGCGGAAACCGACGCGGAGGGGTTCAAGCTTCAACACAC 931  
 Qy 274 AlaAspPheAsnIleValGlyAlaAlaAlaLysGlyGlyGlyAlaAlaGlyAsp 293  
 Db 932 ACCGACTTCTACTCCATGTGTGGCGCGCAGCTCCAACTTCGCGCGCGCGCGCGCGCG 991  
 Qy 294 GluGluLysGlyAlaCys----- 299  
 Db 992 GTCCGACCG 1051  
 Qy 300 -----GlyGlyGlyGlyGlyHis-----SerProGlnPro 310  
 Db 1052 TACCGCTCCCGCGCTCGAATGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1111  
 Qy 311 GlnAlaValAlaValPro-----AlaLysArg 319  
 Db 1112 GCCGTGTGTCGGGCG 1171  
 Qy 320 LysAspLeuHisMetLeuValTrpSerSerSerAlaSerProValSerGluArgAlaAla 339  
 Db 1172 GAGGACCTCCACATGTTCTGCTGAGCTCCAGCGCGCTCGCGCGCGCGCGCGCGCGCG 1222  
 Qy 340 ValHisValPheGlyAlaGlyAlaAspHisAlaAspValLeuAla----- 355

Db 1223 -----GTCTTCGGCG 1276  
 Qy 356 -----LysGlyAlaGlnAlaTyArgGlyTy---GlyArgAspAspTyArgSer 371  
 Db 1277 CGCAAAATGATGAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1336  
 Qy 372 ArgThrLysAsnGlySerGlyAlaAspLysGlyGlyProThrLeuSerLysLeuGly 391  
 Db 1337 GCGAACACAGG-----GCGCTCATGACAGGAGC-----GCGGAGGCGAGG 1375  
 Qy 392 SerAsnSerThrAlaGlnLeuTyProLysAspAspGlyGluGlyArgAlaAlaVal 411  
 Db 1376 GACGAGAGCG 1435  
 Qy 412 AlaMetProProAlaSerValMetThrArgLeuLeuLeuMetValTrpArgLysLeu 431  
 Db 1436 GCGATGCG 1495  
 Qy 432 IleArgAsnProAsnThrTyArgSerLeuLeuGlyValTrpSerLeuValSerTy 451  
 Db 1496 ATCCGCAACCGCAACCTACTCCAGCCTCATCGGCTCATCTGGTCCCTCGTCTGCTTC 1555  
 Qy 452 ArgTrpGlyIleGluMetProAlaIleAlaAlaArgSerIleSerIleLeuSerAspAla 471  
 Db 1556 AGGTGGAACCTTCAGATCGCGCGCATCGTCTGAAATCCATCTCGATCTCTGCGGCG 1615  
 Qy 472 GlyLeuGlyMetAlaMetPheSerLeuGlyLeuPheMetAlaLeuGlnProArgIleIle 491  
 Db 1616 GGGCTCGGATGCGCATGTTTCAGTCTCGGTCCTTCATGCGCTGCGCGCGCATCATC 1675  
 Qy 492 AlaCysGlyAsnLysLeuAlaAlaIleAlaMetGlyValArgPheValAlaGlyProAla 511  
 Db 1676 GCGTGGGGAACAGGTGGCGCATCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1735  
 Qy 512 ValMetAlaAlaAlaSerIleAlaValGlyLeuArgGlyValLeuLeuHisIleAlaIle 531  
 Db 1736 GTGATGCGCGCGCGCTCTTCGCGCGTGGACTCCGCGGACGCTCTCTGCGCGCGCG 1795  
 Qy 532 ValGlnAlaAlaLeuProGlnGlyIleValProPheValPheAlaLysGlyTyGlyVal 551  
 Db 1796 GTCAGGACGCTCTGCCCGGAGCATGTCCCTTCGTCTTCGCGCGGAGGTACAGCGTG 1855  
 Qy 552 HisProAspIleLeuSerThrAla-----TyGly-----ProlLeuThr 564  
 Db 1856 CACCTAGCATTTCTCAGCAGCTGTCACTTTGGCATGTCTATCGCTTGCCTATCACC 1915

RESULT 5  
 ADA70668  
 ID ADA70668 standard; DNA; 1827 BP.  
 XX ADA70668;  
 AC ADA70668;  
 XX 20-NOV-2003 (first entry)  
 DT  
 XX  
 XX Rice gene, SEQ ID 3991.  
 DE  
 XX Plant; bacterial infection; fungal infection; viral infection; rice;  
 KW gene; ds.  
 XX  
 OS Oryza sativa.  
 XX  
 XX WO2003000898-A1.  
 PN  
 XX  
 XX 03-JAN-2003.  
 PD  
 XX  
 XX 22-JUN-2001; 2001WO-IB001105.  
 PF  
 XX  
 XX 22-JUN-2001; 2001WO-IB001105.  
 PR  
 XX  
 XX (SYGN ) SYNGENTA PARTICIPATIONS AG.  
 PA  
 XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
 PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;

XX WPI; 2003-175290/17.  
 XX Identifying at least one gene involved in plant resistance or response to  
 PT pathogenic infection for conferring resistance or tolerance to a plant to  
 PT bacterial, fungal or viral infection by determining or detecting plant  
 PT gene expression.

XX Claim 6; SEQ ID NO 3991; 899bp; English.

XX The present invention relates to a method (M1) for identifying genes  
 CC involved in plant resistance or response to pathogenic infection. M1  
 CC comprises identifying a gene whose expression is significantly altered in  
 CC the incompatible interaction of plant gene expression relative to  
 CC expression of the gene in an unfected plant, in a mutant plant that  
 CC does not express a gene associated with response to pathogenic infection,  
 CC or in a corresponding incompatible or compatible interaction. (M1) is  
 CC useful for conferring resistance to resistance or tolerance to a plant to  
 CC bacterial, fungal or viral infection. The present sequence was used to  
 CC illustrate the invention.

XX SQ Sequence 1827 BP; 284 A; 638 C; 571 G; 332 T; 0 U; 2 Other;

#### Alignment Scores:

Pred. No.: 2,81e-106 Length: 1827  
 Score: 1647,50 Matches: 374  
 Percent Similarity: 66.09% Conservative: 45  
 Best Local Similarity: 58.99% Mismatches: 100  
 Query Match: 56.62% Indels: 115  
 DB: 7 Gaps: 18

US-10-030-884-14 (1-573) x ADA70668 (1-1827)

QY 11 LeuThrAlaValProLeuTyrValAlaMetThrLeuAlaTyrGlySerValArgTyr 30  
 DB :::::  
 QY 1 ATGGCGGGGGTGGTGGCGGCTAGCTGGCGATGTCTCTGGCGTACGGGTGGTGGTGG 60  
 DB :::::  
 QY 31 TrpArgIlePheThrProAspGlnCysSerGlyIleAsnArgPheValAlaLeuPheAla 50  
 DB :::::  
 QY 61 TGGGCAITCTTACCGCCGACAGTGTCTCGGATCAACCGTCTGCTGCCATCTTCCGC 120  
 DB :::::  
 QY 51 ValProLeuLeuSerPheHisPheIleSerThrAsnAspPropheAlaMetAsnLeuArg 70  
 DB :::::  
 QY 121 GTGGCGCTCTTCCACTTCATCTCCACCAACACCGCTACGCCATGAACCTCCGC 180  
 QY 71 PheLeuAlaAlaAspThrLeuGlnLysValAlaValLeuAlaLeuAlaLeuAlaSer 90  
 DB :::::  
 QY 181 TTCTTGGCGGGGACACGCTGAGAGAGTGTCTGCTGGCGGGCTGCGCGCGTGTGTCG 240  
 QY 91 ArgGlyLeuSerSerProArgAlaLeuGlyLeuAspTrpSerIleThrLeuPheSerLeu 110  
 DB :::::  
 QY 241 CGCTTCCCTCGCGACCGCGCGCGCTGGACTGGTCCATCAGCTCTTCTCCCTC 300  
 QY 111 SerThrLeuProAsnThrLeuValMetGlyIleProLeuLeuArgGlyMetTyrGlyAla 130  
 DB :::::  
 QY 301 TCCACGCTGCCCAACACGCTGTCTATGGGATCCCGCTGCTGATCGCCATGTACGGGCA 360  
 QY 131 SerSerAlaGlyThrLeuMetValGlnValValValLeuGlnCysIleIleTyrThr 150  
 DB :::::  
 QY 361 TACTTC---GGCTGCTCATGTCTCAGATCGTGTGCTCTCAGTGCATCATCTGTACAGC 417  
 QY 151 LeuMetLeuPheLeuPheGluTyrArgAlaAlaArgAlaLeuValLeuAspGlnPhePro 170  
 DB :::::  
 QY 418 CTGATGCTCTTCTCTTGGAGTTCGCGCGCGCGGATGTGATGCGCGACAGTTCGCG 477  
 QY 171 AspGlyAlaAlaAlaSerIleValSerPheArgValAspSerAspValValSerLeuAla 190  
 DB :::::  
 QY 478 GAC---ACGGCGGCGTCCATCTGTCTTGCACGCTCGACCGGACGTGTGTCTGCTGGAG 534  
 QY 191 ArgGlyAspValGluLeuGluAlaGluProAspGlyValAlaGlyAlaValSer 210  
 DB :::::  
 QY 535 GCGCGCCACCGGACGAGCGAGCGCGAG-----GTGGCGGCG----- 570

QY 211 SerArgGlyGlyAspAlaGlyArgValArgValThrValArgLysSerThrSerSerArg 230  
 DB :::::  
 QY 571 -----GACGGCGGCTGCACGTCCACGTGCGCGGTCTCTCGGTGTCGGG 615  
 QY 231 SerGluAlaAlaCysSerHisSerHisSerGlnThrMetGlnProArgValSerAsnLeu 250  
 DB :::::  
 QY 616 -----CGTCTGCTGGTTCACGCGCGGCGGTGCAACCTG 651  
 QY 251 SerGlyValGluIleTyrSerLeuGlnSerSerArgAsnProThrProArgGlySerSer 270  
 DB :::::  
 QY 652 ACGGAGCGAGATCTACTCTAGCTGTGCGGACCCCAACCGCGCGGCTCCAAAC 711  
 QY 271 PheAsnHisAlaAspPhePheAsnIleValGly-----Ala 282  
 DB :::::  
 QY 712 TTCACACGCGGATCTTCTGCCATGTGCGGGGCGGCCACCGCCCGCGCGCGT 771  
 QY 283 AlaAlaLysGlyGlyGlyAlaAlaGlyAsp----- 293  
 DB :::::  
 QY 772 GCGTGGCGGCTCGAGCTTCGCGCTCCGAGCTTACTCTGCTCAATCTGCGGGGAC 831  
 QY 294 -----GluGluLysGlyAlaCysGlyGlyGlyGlyGlyGly 305  
 DB :::::  
 QY 832 CCAACCCCGAGGAGTCCAACTTCGACGAGCACTCGGCA----- 870  
 QY 306 HisSerProGlnProGlnAlaValAlaValProAla-----LysArgLysAspLeu 322  
 DB :::::  
 QY 871 CGGCGCGGAAACCCGCGCACCCACGCGGCGCTCAACACGATGCCAAGGAGCTC 930  
 QY 323 HisMetLeuValTrpSerSerSerAlaSerProValSerGluArgAlaAlaValHisVal 342  
 DB :::::  
 QY 931 CACATGTTCTGTGGAGCTCGAGCGCTCTCCGCTCTCAGAGTCAGCGCGCTGCTGTG 990  
 QY 343 Phe---GlyAlaGlyGlyAlaAspHisAlaAspValLeuAlaLys----- 356  
 DB :::::  
 QY 991 TTCTGTTGGCGGGCGCGCGCGCTCTCGAGCTCGGCGCCCAAGAAATCCACATGTC 1050  
 QY 357 -----GlyAlaGlnAlaTyrAspGluTyrGlyArg 366  
 DB :::::  
 QY 1051 ATCCCGCGACCTGCGCGCAGAACACCGCTCAGGCAAGACGACGAGAGTACGCGC--- 1107  
 QY 367 AspAspTyrSerSerArgThrLysAsnGlySerGlyGlyAlaAspLys----- 382  
 DB :::::  
 QY 1108 -----GCACTGGCATTTGGTGGCGCGCGCGGAGAGACTTCAGCTTCGGA 1155  
 QY 383 GlyGlyProThr----- 386  
 DB :::::  
 QY 1156 GCGCGCAACGCTGAGCGCGCGGCGAGGAGTAGAGGAGGCGGCTTGTCTGACGGG 1215  
 QY 387 LeuSerLysLeuGlySerAsnSerThrAlaGlnLeuTyrProLys-----AspAsp 403  
 DB :::::  
 QY 1216 CTGACGAGATGGGTGCGAGCTCGACGCGGAGCTGCACCCGAGGTCTCGACGTGAC 1275  
 QY 404 GlyGluGlyArgAlaAlaAlaValAla-----MetProProAlaSer 417  
 DB :::::  
 QY 1276 GGAACCGACCGCGCGCGCGCGCGCGCGAGTACCAATGCGCGCGGCGAGC 1335  
 QY 418 ValMetThrArgLeuIleLeuMetValTrpArgLysLeuIleArgAsnProAsnThr 437  
 DB :::::  
 QY 1336 GTGATGACACGCTCATCTCTATAATGGTGGCGGCAAGCTCATTCGCAACCCCAACT 1395  
 QY 438 TyrSerSerLeuIleGlyValValTrpSerLeuValSerTyrArgTrpGlyIleGluMet 457  
 DB :::::  
 QY 1396 TACTCCAGCTCTCGGCTCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1455  
 QY 458 ProAlaIleAlaAlaArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMet 477  
 DB :::::  
 QY 1456 CCAGCAATGTCGAGAGTCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1515  
 QY 478 PheSerLeuGlyLeuPheMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeu 497  
 DB :::::  
 QY 1516 TTTAGCTGGGATTTGTTTCATGGCGCTGCGAGCCGAGCATCTATCCGCTGTGGCAATCAGCC 1575  
 QY 498 AlaAlaIleAlaMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaAlaSer 517

1576 GCGTGTCTCCATGGCGCTCGCTTCTTCGCGGGCCCTGCGTGCATGCGCGCGTCA 1635  
 518 ILeAlaValGlyLeuArgGlyValLeuLeuHisIleAlaIleValGlnAlaAlaLeuPro 537  
 1636 ATCGCATCGGACTCGCGGGAGCGTCTCTGCGAGTGCAGTGTTCAGCGGCGCTTACCA 1695  
 538 GlnGlyIleValProPheValPheAlaLysGlnGlyValHisProAspIleLeuSer 557  
 1696 CAAGGAGTGTGCTTTGTTTGGCAAGAATAACAATGTCCACCGCGCCATCCTGAGC 1755  
 558 ThrAla-----TyrGly-----ProIleThr 564  
 1756 ACAGCGTAAATTTTGGCATGTCTAATAGCTCTTCCATACACA 1797

RESULT 6  
 ID AAA94729 standard; DNA; 2549 BP.  
 AC AAA94729;  
 XX  
 DT 02-FEB-2001 (first entry)  
 XX Soybean auxin transport protein clone sfl1.pk131.g9 DNA sequence.  
 DE  
 XX Auxin transport protein; soybean; root development; gene mapping;  
 KW  
 KW plant breeding; herbicide; ss.  
 XX  
 OS Glycine max.  
 XX  
 PN WO200068389-A2.  
 XX  
 PD 16-NOV-2000.  
 XX  
 PF 03-MAY-2000; 2000WO-US012061.  
 XX  
 PR 07-MAY-1999; 99US-0133040P.  
 XX  
 PA (DUPO) DU PONT DE NEMOURS & CO E I.  
 PA (PION-) PIONEER HI-BRED INT INC.  
 XX  
 PI Orozco EM, Weng Z, Bruce WB, Cahoon RE, Tao Y;  
 XX  
 DR WPI; 2000-687647/67.  
 DR P-PSDB; AAB26942.  
 XX

PT New nucleic acid sequences encoding new auxin transport proteins, useful  
 for modulating root growth of plants and to screen for herbicides.  
 XX  
 PS Claim 2; Page 74-75; 94pp; English.  
 XX  
 CC Auxins are plant hormones that influence plant behaviour and development  
 CC e.g. vascular tissue differentiation, apical development, tropic  
 CC responses and organ (e.g. flower, leaf) development. The present  
 CC invention relates to corn auxin transport protein coding sequences and  
 CC proteins. The present sequence is one such coding sequence. This sequence  
 CC may be used to modulate root development, e.g. to produce a more robust  
 CC root system, alter root angle or redirect root growth. Also, the present  
 CC sequence may be useful for gene mapping (e.g. for plant breeding) and to  
 CC identify loss of function mutants. The protein encoded by the present  
 CC sequence may be useful for raising specific antibodies, for the detection  
 CC of auxin transport proteins and to design and/or identify specific  
 CC inhibitors of auxin transport proteins, potentially useful as herbicides

XX  
 SQ Sequence 2549 BP; 736 A; 572 C; 518 G; 723 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 8,55e-106 Length: 2549  
 Score: 1643.00 Matches: 351  
 Percent Similarity: 66.51% Conservative: 58  
 Best Local Similarity: 57.30% Mismatches: 111  
 Query Match: 56.46% Indels: 100  
 3 Gaps: 17  
 DB:

US-10-030-884-14 (1-573) x AAA94729 (1-2549)  
 Qy 1 MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValProLeuTyrValAla 20  
 Db 309 ATGATCACCTTACAGACATCTTACCATGTGATGCAATGGTGCACCTCTATGTGGCC 368  
 Qy 21 MetThrIleuAlaTyrGlySerValArgTyrPArgIlePheThrProAspGlnCysSer 40  
 Db 369 ATGATACCTATGCGCTCAGTGAAGTGGTGAAGATTTTCTCCCTGATCAATGCTCT 428  
 Qy 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
 Db 429 GGCATCAACCGTTTGTGGCCTCTTTGGCAGTGCCTCTCTCTCTCCCTCCACTTCATAGCC 488  
 Qy 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrIleuGlnLysVal 80  
 Db 489 TCCACCAACCCCTTATGAGATGAACCTGAGGTTCCTAGCTGCTGACACCCCTTCAAAAGATC 548  
 Qy 81 AlaValLeuAlaLeuAla-----LeuAlaSerArgGlyLeuSerSerPro 96  
 Db 549 ATAATACTAGTCTCTCTTGGCAGTTCGGAGCAACATCACCAAAAGGGGTGT----- 599  
 Qy 97 ArgAlaLeuGlyLeuAspTyrPserIleThrLeuPheSerLeuSerThrLeuProAsnThr 116  
 Db 600 -----TTGGAATGGGCATAACCTTCTCTCTCTCTCCACCTTCCCAACACT 647  
 Qy 117 LeuValMetGlyIleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeu 136  
 Db 648 TTGTTATGGCATCCCTTTGCTCAAGAGGATGATGGTCACTTCA-----GGAGGCTC 704  
 Qy 137 MetValGlnValValLeuGlnCysIleIleTyrThrLeuMetLeuPheLeuPhe 156  
 Db 705 ATGGTGCAATTTGGTCTCTCCAGTGCATCATTTGGTACACCTTGATGCTCTCTTGTGTT 764  
 Qy 157 GluTyrArgAlaAlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAsp 176  
 Db 765 GAGTTTAGAGTGCAGATGCTCATCTCTGAGCAGTTCCTCTGAC---ACTGTGCTCTCC 821  
 Qy 177 IleValSerPheArgValAspSerAspValValSerLeu---AlaArgGlyAspValGlu 195  
 Db 822 ATTGTCTCCATCCATGTGGACTCTGATGTCATGTCATGGATGGAAGACCAACACTTGAG 881  
 Qy 196 LeuGluAlaGluProAspGlyValAlaGlyAlaValSerSerArgGlyGlyAsp 215  
 Db 882 ACTGAAGCTGAGATCAAG-----GAA 902  
 Qy 216 AlaGlyArgValArgValThrValArgLysSerThrSerSerArgSerGluAlaAlaCys 235  
 Db 903 GATGGTAACTCCATGTCATCTGTGAGGAATCCAATGCTTCAAGATCAGACATCTTCTCT 962  
 Qy 236 SerHisSerHisSer---GlnThrMetGlnProArgValSerAsnLeuSerGlyValGlu 254  
 Db 963 AGAAGGTCTCAGGGTCTCTCTTCCACCATCCACGCCCTTCCAACTTACCAATGCTGAG 1022  
 Qy 255 IleTyrSerLeuGlnSerSerArgAsnProThrProArgGlySerSerPheAsnHisAla 274  
 Db 1023 ATATCTCTTTGCANCTCTTAGAACCTTACGCCGAGAGGCTTAGTTTCAACCACT 1082  
 Qy 275 AspPhePheAsnLeuValGlyAlaAlaLysGlyGlyGlyAlaAlaGlyAsp--- 293  
 Db 1083 GATTTCTACTCCATGATGGTGTGGTGGCAGGAACCTCAAACTTTGGTGGCTCTGATGTT 1142  
 Qy 294 -----GluGluLysGly 297  
 Db 1143 TATGGCTTTTCAAGTTCAGGCGCCCACTCCAGGCCCTTCTAACTATGATGATGATGGT 1202  
 Qy 298 Ala-----CysGlyGlyGlyGlyGlyGlyHisSerProGlnProGln 311  
 Db 1203 GGGAAAGCAAAAGTTTTCATTACCATGCTGCTGGTGGAACTGGGCACTACCTCGCACCAC 1262  
 Qy 312 -----AlaValAlaValProAlaLys 318













Db 1054 AAGGAGGTCGATGGCGTGGCTCGCGCGCAAGCGGACGGGTGGAGAGGACGAC 1113  
 Qy 369 TyrSerSer-----ArgThrLysSerGlySerGly-AlaAspLysGlyG1 384  
 Db 1114 TTCAGCTT-CGGGAACAGAGCGCTCGCGGAGGAGCGCGGAGCGCGGACGAGAAG 1172  
 Qy 384 yProThrLeuSerLysLeuGlySerAsnSerThrAlaGlnLeuTyrProLysAspAspG1 404  
 Db 1173 CGTGGCGCGGGGTGGGTGGAGCATGGCAAGCTGGCTTGACGCG----- 1221  
 Qy 404 yGluGlyArgAlaAlaAlaValMetProProAlaSerValMetThrArgLeuLeuLe 424  
 Db 1222 -----GGCGGAGCGGGATGCCCGGACGAGCGGTGATGACGGCGCTCATCT 1268  
 Qy 424 uileMetValTyrArgLysLeuLeuArgAsnProAsnThrTyrSerSerLeuLeuGlyVa 444  
 Db 1269 GATCATGTGTGGCGAGCATCATCGCAACCCCAACACCTACTCCAGCCTCATCGGCT 1328  
 Qy 444 lValTyrSerLeuValSerTyrArgTyrGlyLeuMetProAlaLeuLeuLeuAlaArgSe 464  
 Db 1329 CATCTGGTGGCTGTCTGCTGCTGGTGAACCTTCAGATGGCGGCGATCATCTGAATC 1388  
 Qy 464 rIleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeuGlyLeuPheMe 484  
 Db 1389 CATCTTCATCTCTCCGAGCGCGCTCGGATGGCCATGTTACGCTCGGCGCTTCAT 1448  
 Qy 484 tAlaLeuGlnProArgIleLeuAlaCysGlyAsnLysLeuAlaAlaLeuMetGlyVa 504  
 Db 1449 GCGCTGCGAGCGCGATCATCGCTGCGGGAACAAGGTGGCGAGTTCGGCATGGCGGT 1508  
 Qy 504 lArgPheValAlaGlyProAlaValMetAlaAlaSerIleAlaValGlyLeuArgG1 524  
 Db 1509 GCGGTTCCTGACCGCGCGCGCGCTCATGGCGCGCGCTCCATCGCGCGCTCGCGG 1568  
 Qy 524 yValLeuLeuHisIleAlaIleValGlnAlaAlaLeuProGlnGlyIleValProPheVa 544  
 Db 1569 CACGCTCTCCAGCTGCCATCGTCGAGCGAGCGGTTCGCMCAGCGCATTTGTCCTTSGT 1628  
 Qy 544 lPheAlaLysGluTyrGlyValHisProAspIleLeuSerThrAla 559  
 Db 1629 CTTCCGCAAGGAGTACAGCTGCACCCGACATTCACGACGCGG 1674

RESULT 10  
 BAA94733  
 ID AAA94733 standard; DNA; 2293 BP.  
 AC AAA94733;  
 DT 02-FEB-2001 (first entry)  
 DE Wheat auxin transport protein clone wdk1c.pk008.g1 DNA sequence.  
 DE Auxin transport protein; wheat; root development; gene mapping;  
 KW plant breeding; herbicide; ss.  
 OS Triticum aestivum.  
 XX WO200068389-A2.  
 XX 16-NOV-2000.  
 XX 03-MAY-2000; 2000WC-US012061.  
 XX 07-MAY-1999; 99US-0133040P.  
 XX (DUPO) DU PONT DE NEMOURS & CO E I.  
 XX (PION-) PIONEER HI-BRED INT INC.  
 XX Orozco EM, Weng Z, Bruce WB, Cahoon RE, Tao Y;  
 XX WPI: 2000-587647/67.  
 XX P-PSDB; AAB26946.  
 XX

PT New nucleic acid sequences encoding new auxin transport proteins, useful  
 for modulating root growth of plants and to screen for herbicides.  
 XX Claim 2; Page 84; 94pp; English.  
 CC Auxins are plant hormones that influence plant behaviour and development  
 e.g. vascular tissue differentiation, apical development, tropic  
 responses and organ (e.g. flower, leaf) development. The present  
 CC invention relates to corn auxin transport protein coding sequences and  
 CC proteins. The present sequence is one such coding sequence. This sequence  
 CC may be used to modulate root development, e.g. to produce a more robust  
 CC root system, alter root angle or redirect root growth. Also, the present  
 CC sequence may be useful for gene mapping (e.g. for plant breeding) and to  
 CC identify loss of function mutants. The protein encoded by the present  
 CC sequence may be useful for raising specific antibodies, for the detection  
 CC of auxin transport proteins and to design and/or identify specific  
 CC inhibitors of auxin transport proteins, potentially useful as herbicides  
 XX

SQ Sequence 2293 BP; 462 A; 713 C; 683 G; 435 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 1.75e-100 Length: 2293  
 Score: 1566.50 Matches: 364  
 Percent Similarity: 65.26% Conservative: 53  
 Best Local Similarity: 56.96% Mismatches: 132  
 Query Match: 53.83% Indels: 91  
 DB: 3 Gaps: 17

US-10-030-884-14 (1-573) x AAA94733 (1-2293)

Qy 1 MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValProLeuTyrValAla 20  
 Db 74 ATGATCACCAGGAGGAGCATCTACGACGTCGTGGCGCGGTGGTGGCTGTACGTGGCC 133  
 Qy 21 MetThrLeuAlaTyrGlySerValArgTyrTyrArgIlePheThrProAspGlnCysSer 40  
 Db 134 ATGTTTCATGTGGGTACGGGTGCGGTGCGGTGGGCGATCTTCACGCGGACCATGCTCG 193  
 Qy 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
 Db 194 GGATCAACCGCTTCGTGGCGGTCTTCGGCGGTGCGCTCTCTCTCTCTCTCTCTCTCT 253  
 Qy 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80  
 Db 254 ACCACGAGCCCTACGCCATGGACTACCGCTTCTGGCGCGGACTCGGTGACGAGAGTC 313  
 Qy 81 AlaValLeuAlaLeuAlaLeuAlaSerArgGlyLeuSerSerProArgAlaLeuGly 100  
 Db 314 GTCATCTCGCGCGCTTCGCGGTGGGACAAAGCTGCTCTCCGCTACCGGTGCGCGCGC 373  
 Qy 101 -----LeuAspTyrSerIleThrLeuPheSerLeuSer 111  
 Db 374 GGCACGAGCGCGCGAGGCTCGTGGTGGAGTGGACCATCAGCTCTCTCTCTCTCTCT 433  
 Qy 112 ThrLeuProAsnThrLeuValMetGlyIleProLeuLeuArgGlyMetTyrGlyAlaSer 131  
 Db 434 ACGTGTCCCAACACGCTGGTGGATCCGCTCTCGCGCGCATGTACGCGGACTTC 493  
 Qy 132 SerAlaGlyThrLeuMetValGlnValValValLeuGlnCysIleIleTyrThrLeu 151  
 Db 494 TCG---GGGTGCTCATGTGTGAGATCTGGTGTGTCGAGCGGTCTCTGGTACACGCTC 550  
 Qy 152 MetLeuPheLeuPheGluTyrArgAlaAlaArgAlaLeuValLeuAspGlnPheProAsp 171  
 Db 551 ATGCTCTTCTCTTCGAGTACCGCGCGCGGACAGCGCTCATCTCCGAGCAGTTCGCGGCC 610  
 Qy 172 GlyAlaAlaAlaSerIleValSerPheArgValAspSerAspValValSerLeuAlaArg 191  
 Db 611 GACGTGCGCGCGCGAGCATCGCTCTTCGCGGTGAGTCCGAGCTCTCTCTCTCTCTCT 667  
 Qy 192 GlyAspValGluLeuGluAlaGluProAspGlyValAlaGlyAlaGlyAlaValSerSer 211  
 Db 668 GGGCGCGAGCGCTGCGACGCGGAC-----GCC 694

513	533	1758	533	1828	533	1888	RESULT 11		AAC44722		AAC44722 standard; DNA, 1923 BP.		AAC44722;		18-OCT-2000 (first entry)		Arabidopsis thaliana DNA fragment SEQ ID NO: 43888.		Hybridisation assay; Genetic mapping; gene expression control;		protein identification; signal transduction pathway; metabolic pathway;		promoter; termination sequence; ss.		Arabidopsis thaliana.		EP1033405-A2.		06-SEP-2000.		25-FEB-2000; 2000EP-00301439.		25-FEB-1999; 99US-0121825P.		05-MAR-1999; 99US-0123180P.		09-MAR-1999; 99US-0123548P.		23-MAR-1999; 99US-0125788P.		25-MAR-1999; 99US-0126264P.		29-MAR-1999; 99US-0126785P.		01-APR-1999; 99US-0127452P.		06-APR-1999; 99US-0128234P.		08-APR-1999; 99US-0128714P.		16-APR-1999; 99US-0129845P.		19-APR-1999; 99US-0130077P.		21-APR-1999; 99US-0130449P.		23-APR-1999; 99US-0130891P.		28-APR-1999; 99US-0131449P.		30-APR-1999; 99US-0132048P.		30-APR-1999; 99US-0132407P.		04-MAY-1999; 99US-0132484P.		05-MAY-1999; 99US-0132485P.		06-MAY-1999; 99US-0132486P.		06-MAY-1999; 99US-0132487P.		07-MAY-1999; 99US-0132863P.		11-MAY-1999; 99US-0134256P.		14-MAY-1999; 99US-0134218P.		14-MAY-1999; 99US-0134219P.		14-MAY-1999; 99US-0134221P.		18-MAY-1999; 99US-0134370P.		19-MAY-1999; 99US-0134768P.		20-MAY-1999; 99US-0134941P.		21-MAY-1999; 99US-0135353P.		24-MAY-1999; 99US-0135629P.		25-MAY-1999; 99US-0136021P.		27-MAY-1999; 99US-0136392P.		28-MAY-1999; 99US-0136782P.		01-JUN-1999; 99US-0137222P.		03-JUN-1999; 99US-0137528P.		04-JUN-1999; 99US-0137502P.		07-JUN-1999; 99US-0137724P.		513		533		1758		533		1828		533		1888		RESULT 11		AAC44722		AAC44722 standard; DNA, 1923 BP.		AAC44722;		18-OCT-2000 (first entry)		Arabidopsis thaliana DNA fragment SEQ ID NO: 43888.		Hybridisation assay; Genetic mapping; gene expression control;		protein identification; signal transduction pathway; metabolic pathway;		promoter; termination sequence; ss.		Arabidopsis thaliana.		EP1033405-A2.		06-SEP-2000.		25-FEB-2000; 2000EP-00301439.		25-FEB-1999; 99US-0121825P.		05-MAR-1999; 99US-0123180P.		09-MAR-1999; 99US-0123548P.		23-MAR-1999; 99US-0125788P.		25-MAR-1999; 99US-0126264P.		29-MAR-1999; 99US-0126785P.		01-APR-1999; 99US-0127452P.		06-APR-1999; 99US-0128234P.		08-APR-1999; 99US-0128714P.		16-APR-1999; 99US-0129845P.		19-APR-1999; 99US-0130077P.		21-APR-1999; 99US-0130449P.		23-APR-1999; 99US-0130891P.		28-APR-1999; 99US-0131449P.		30-APR-1999; 99US-0132048P.		30-APR-1999; 99US-0132407P.		04-MAY-1999; 99US-0132484P.		05-MAY-1999; 99US-0132485P.		06-MAY-1999; 99US-0132486P.		06-MAY-1999; 99US-0132487P.		07-MAY-1999; 99US-0132863P.		11-MAY-1999; 99US-0134256P.		14-MAY-1999; 99US-0134218P.		14-MAY-1999; 99US-0134219P.		14-MAY-1999; 99US-0134221P.		18-MAY-1999; 99US-0134370P.		19-MAY-1999; 99US-0134768P.		20-MAY-1999; 99US-0134941P.		21-MAY-1999; 99US-0135353P.		24-MAY-1999; 99US-0135629P.		25-MAY-1999; 99US-0136021P.		27-MAY-1999; 99US-0136392P.		28-MAY-1999; 99US-0136782P.		01-JUN-1999; 99US-0137222P.		03-JUN-1999; 99US-0137528P.		04-JUN-1999; 99US-0137502P.		07-JUN-1999; 99US-0137724P.		513		533		1758		533		1828		533		1888		RESULT 11		AAC44722		AAC44722 standard; DNA, 1923 BP.		AAC44722;		18-OCT-2000 (first entry)		Arabidopsis thaliana DNA fragment SEQ ID NO: 43888.		Hybridisation assay; Genetic mapping; gene expression control;		protein identification; signal transduction pathway; metabolic pathway;		promoter; termination sequence; ss.		Arabidopsis thaliana.		EP1033405-A2.		06-SEP-2000.		25-FEB-2000; 2000EP-00301439.		25-FEB-1999; 99US-0121825P.		05-MAR-1999; 99US-0123180P.		09-MAR-1999; 99US-0123548P.		23-MAR-1999; 99US-0125788P.		25-MAR-1999; 99US-0126264P.		29-MAR-1999; 99US-0126785P.		01-APR-1999; 99US-0127452P.		06-APR-1999; 99US-0128234P.		08-APR-1999; 99US-0128714P.		16-APR-1999; 99US-0129845P.		19-APR-1999; 99US-0130077P.		21-APR-1999; 99US-0130449P.		23-APR-1999; 99US-0130891P.		28-APR-1999; 99US-0131449P.		30-APR-1999; 99US-0132048P.		30-APR-1999; 99US-0132407P.		04-MAY-1999; 99US-0132484P.		05-MAY-1999; 99US-0132485P.		06-MAY-1999; 99US-0132486P.		06-MAY-1999; 99US-0132487P.		07-MAY-1999; 99US-0132863P.		11-MAY-1999; 99US-0134256P.		14-MAY-1999; 99US-0134218P.		14-MAY-1999; 99US-0134219P.		14-MAY-1999; 99US-0134221P.		18-MAY-1999; 99US-0134370P.		1	
-----	-----	------	-----	------	-----	------	-----------	--	----------	--	----------------------------------	--	-----------	--	---------------------------	--	---	--	--	--	---	--	-------------------------------------	--	-----------------------	--	---------------	--	--------------	--	-------------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----	--	-----	--	------	--	-----	--	------	--	-----	--	------	--	-----------	--	----------	--	----------------------------------	--	-----------	--	---------------------------	--	---	--	--	--	---	--	-------------------------------------	--	-----------------------	--	---------------	--	--------------	--	-------------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----	--	-----	--	------	--	-----	--	------	--	-----	--	------	--	-----------	--	----------	--	----------------------------------	--	-----------	--	---------------------------	--	---	--	--	--	---	--	-------------------------------------	--	-----------------------	--	---------------	--	--------------	--	-------------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	-----------------------------	--	---	--

PR	08-JUN-1999;	99US-0138094P.	PR	09-AUG-1999;	99US-0147493P.
PR	10-JUN-1999;	99US-0138540P.	PR	09-AUG-1999;	99US-0147935P.
PR	10-JUN-1999;	99US-0138847P.	PR	10-AUG-1999;	99US-0148171P.
PR	14-JUN-1999;	99US-0139111P.	PR	11-AUG-1999;	99US-0148319P.
PR	16-JUN-1999;	99US-0139452P.	PR	12-AUG-1999;	99US-0148341P.
PR	16-JUN-1999;	99US-0139453P.	PR	13-AUG-1999;	99US-0148565P.
PR	17-JUN-1999;	99US-0139492P.	PR	13-AUG-1999;	99US-0148684P.
PR	18-JUN-1999;	99US-0139454P.	PR	15-AUG-1999;	99US-0149368P.
PR	18-JUN-1999;	99US-0139455P.	PR	17-AUG-1999;	99US-0149175P.
PR	18-JUN-1999;	99US-0139456P.	PR	18-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	99US-0139457P.	PR	20-AUG-1999;	99US-0149722P.
PR	18-JUN-1999;	99US-0139458P.	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	99US-0139459P.	PR	22-AUG-1999;	99US-0149929P.
PR	18-JUN-1999;	99US-0139460P.	PR	23-AUG-1999;	99US-0149902P.
PR	18-JUN-1999;	99US-0139461P.	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999;	99US-0139462P.	PR	25-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	99US-0139463P.	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	99US-0139750P.	PR	27-AUG-1999;	99US-0151065P.
PR	18-JUN-1999;	99US-0139763P.	PR	27-AUG-1999;	99US-0151066P.
PR	21-JUN-1999;	99US-0139817P.	PR	27-AUG-1999;	99US-0151080P.
PR	21-JUN-1999;	99US-0139889P.	PR	30-AUG-1999;	99US-0151303P.
PR	22-JUN-1999;	99US-0140353P.	PR	31-AUG-1999;	99US-0151438P.
PR	23-JUN-1999;	99US-0140354P.	PR	01-SEP-1999;	99US-0151930P.
PR	23-JUN-1999;	99US-0140695P.	PR	07-SEP-1999;	99US-0152363P.
PR	24-JUN-1999;	99US-0140823P.	PR	10-SEP-1999;	99US-0153070P.
PR	28-JUN-1999;	99US-0140991P.	PR	13-SEP-1999;	99US-0153758P.
PR	29-JUN-1999;	99US-0141287P.	PR	15-SEP-1999;	99US-0154018P.
PR	30-JUN-1999;	99US-0141842P.	PR	16-SEP-1999;	99US-0154039P.
PR	01-JUL-1999;	99US-0142154P.	PR	20-SEP-1999;	99US-0154779P.
PR	02-JUL-1999;	99US-0142055P.	PR	22-SEP-1999;	99US-0155139P.
PR	06-JUL-1999;	99US-0142390P.	PR	23-SEP-1999;	99US-0155486P.
PR	08-JUL-1999;	99US-0142803P.	PR	24-SEP-1999;	99US-0155659P.
PR	09-JUL-1999;	99US-0142920P.	PR	25-SEP-1999;	99US-0156458P.
PR	12-JUL-1999;	99US-0142977P.	PR	25-SEP-1999;	99US-0156596P.
PR	13-JUL-1999;	99US-0143542P.	PR	04-OCT-1999;	99US-0157117P.
PR	14-JUL-1999;	99US-0143624P.	PR	05-OCT-1999;	99US-0157753P.
PR	15-JUL-1999;	99US-0144005P.	PR	06-OCT-1999;	99US-0157865P.
PR	16-JUL-1999;	99US-0144085P.	PR	07-OCT-1999;	99US-0158029P.
PR	16-JUL-1999;	99US-0144086P.	PR	08-OCT-1999;	99US-0158232P.
PR	19-JUL-1999;	99US-0144325P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144333P.	PR	14-OCT-1999;	99US-0159329P.
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159330P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159637P.
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159638P.
PR	20-JUL-1999;	99US-0144884P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0144814P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160767P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160768P.
PR	21-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160810P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.
PR	26-JUL-1999;	99US-0145276P.	PR	23-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161406P.
PR	27-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161920P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161922P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161933P.
PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.
PR	04-AUG-1999;	99US-0147302P.			
PR	05-AUG-1999;	99US-0147192P.			
PR	05-AUG-1999;	99US-0147260P.			
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.			

Alignment Scores:  
Pred. No.: 3,75e-100  
Score: 1560.50  
Percent Similarity: 63.87%

Length: 1923  
Matches: 353  
Conservative: 73

Best Local Similarity:	52.92%	Mismatches:	102
Query Match:	53.63%	Indels:	139
DB:	3	Gaps:	20
US-10-030-884-14 (1-573) x AAC44722 (1-1923)			
QY	1	MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValValProLeuTyrValAla	20
DB	1	ATGATCTCAGGACACACCTCTACACGGTCCCTACCGCGGTGATTCCTCTCTACGTGGCC	60
QY	21	MetThrLeuAlaTyrGlySerValArgTyrTrpArgIlePheThrProAspGlnCysSer	40
DB	61	ATGATCTCTCGCTTACGGCTCTGTCGGTGGTGGAAATCTTCTCACCGGACCAATGCTCC	120
QY	41	GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer	60
DB	121	GGAAATCAACCGTTTTCGCGCATCTTTCGCGCTCCCTCTCTCTCTCTCTCTCTCTCTCT	180
QY	61	ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal	80
DB	181	ACCAACAATCTTACGCAATGATCTCCGGTTCATCGCGCGGACACTCTCCAAAATATC	240
QY	81	AlaValLeuAlaLeuLeuAlaLeuAlaSerArgGlyLeuSerSerProArgAlaLeuGly	100
DB	241	ATCATGTGTCTCTTTAGTCTCTATGGCT-----AATTCTACTCGCTCCCGGTAGC	291
QY	101	LeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThrLeuValMetGly	120
DB	292	CTGAGTGGAGCATCAATCTTTTCCCTCTCCACACTTCCCACTCTTCTGTATGGG	351
QY	121	IleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeuMetValGlnVal	140
DB	352	ATTCTCTCTGTATCGCATGTATGGCAATACTCT---GGTTCCCTCATGCTCCAAATC	408
QY	141	ValValLeuGlnCysIleIleTyrThrLeuMetLeuPheLeuPheGluTyrArgAla	160
DB	409	GTGCTCTCCAGGTATCATCTGTCACACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	468
QY	161	AlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSerIleValSerPhe	180
DB	469	GCCAAGATGTCATCATGGGAGCAGTTCCTCTGAG---ACGCTGCTTCCATTTCTTCTTC	525
QY	181	ArgValAspSerAspValValSerLeuAlaArgGlyAsp---ValGluLeuGluAlaGlu	199
DB	526	ANAGTCGAATCCGACGCTGTCGCTCGACGGCCATGATTTCTTGAGACCGATGCGAG	585
QY	200	ProAspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyAlaGlyArgVal	219
DB	586	ATA-----GGTGACGACGGGAAGCTT	606
QY	220	ArgValThrValArgLysSerThrSerSerArgSerGluAlaAlaCysSerHisSerHis	239
DB	607	CACGTCCCGGTGAGAAATCAACGCTTCACGT---CGTTCTGTTCTGC-----	651
QY	240	SerGlnThrMetGlnProAsgValSerAsnLeuSerGlyValGluIleTyrSerLeuGln	259
DB	652	GGCCCGAACATGATCTCACGGCGCTCAAACTCACCGAGCTGAGATTTATAGTCTCAGC	711
QY	260	SerSerArgAsnProThrProArgGlySerSerPheAsnHisAlaAspPheAsnIle	279
DB	712	ACC-----ACTCTAGAGGCTCTAAATTCACCACTCTGATTTTTCACACATG	759
QY	280	ValGly-----Ala	282
DB	760	ATGGGTTTCCCGGTGGTGTCTCTCCAAATTCGGTCCGGCGGATGATGACTCGGTCAA	819
QY	283	AlaAlaLysGly-----GlyGlyGlyAlaAlaGly	292
DB	820	TCAATCAGAGGTCCAACCTCTCGACCTTCAAACCTTCGAGGAGAAATTCGCGCATGGCATCC	879
QY	293	AspGluGluLysGlyAlaCysGlyGlyGlyGlyHisSerProGlnPro-----	310
DB	880	TCCCGAGATTCCGGGTATTATCCCTCGAGGAGGAGCGGGTCTTATCCGGGTCCGAATCCA	939
QY	310	-----	310
DB	940	GAGTTCTTCTTCAACCAACCACTCTACCGCAATAAAGCGTCAATAAAAAACCGGAAGAC	999
QY	311	-----GlnAlaValAlaValProAla-----LysArg	319
DB	1000	GTTAATACGAATCAGCAGACGACTCTTCCAAACGGGGCGCAAGTCAAAACACCCATGAGCC	1059
QY	320	LysAspLeuHisMetLeuValTrpSerSerSerAlaSerProValSerGluArgAlaAla	339
DB	1060	AAGAGTTTCAATGTTCTGAGCTCAACGGGTGTCACCGCTTTCAGACCGGGCGGGT	1119
QY	340	ValHisValPheGlyAlaGlyGlyAlaAspHis-----AlaAspValLeu	354
DB	1120	CTTAACGTTTTCGTC---GGAGCACCTGACAACGATCAAGCGCGGAGAGATCTGCCAAGGT	1176
QY	355	AlaLys-----	356
DB	1177	GCTAAAGAGATCCGATATGTAGTCCAGATCAATCTCAACCGCGGAGACCAAGCTGTA	1236
QY	357	-----GlyAlaGlnAlaTyrAspGluTyrGlyArgAsp	367
DB	1237	GCTCATCCAGCAAGTGGAGATTTTCGGAGGAGAACCACAATTTAGTTTCGCGGAGAAAGAA	1296
QY	368	AspTyrSerSerArgThrLysAsnGlySerGlyGlyAlaAspLysGlyGlyProThrLeu	387
DB	1297	GAAGAAGCAGAGAGACCAAAAGACGCCGAGAAATGTT-----CTA	1335
QY	388	SerLysLeuGlySerAsnSerThrAlaGlnLeuTyrProLysAspAspGlyGluGlyArg	407
DB	1336	AACAACTTGTCTCAATTCACGGCGGCGCTACAATCCAAG---ACAGGTCTAGGAGGA	1392
QY	408	AlaAlaAlaValAla-----MetProProAlaSerValMetThrArgLeuIleLeu	424
DB	1393	GCCGAAGCAAGTCAACGAAAAATATGCTCCGCGAGTGTGATGACAGAGCTGATATCTG	1452
QY	425	IleMetValTrpArgGlyLeuIleArgAsnProAsnThrTyrSerSerLeuIleGlyVal	444
DB	1453	ATAATGGTTTGGAGAAACTCATCAGAAACCCAAACACTTACTTACTTCTCATTTGGACTT	1512
QY	445	ValTrpSerLeuValSerTyrArgTrpGlyIleGluMetProAlaIleIleAlaArgSer	464
DB	1513	ATTGGGCTCTCGTTCGCTTCCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG	1572
QY	465	IleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeuGlyLeuPheMet	484
DB	1573	ATCTCCATCTATCTGATGCTGGTCTTGGAAATGCAATGTTTGGGGTCTTCTCATG	1632
QY	485	AlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAlaAlaIleAlaMetGlyVal	504
DB	1633	CGTTTGCAACCAATTAATCGTTTGGGAATTCAGTGGCAACGTTTGGCATGCGCGTT	1692
QY	505	ArgPheValAlaGlyProAlaValMetAlaAlaSerIleAlaValGlyLeuArgGly	524
DB	1693	AGGTTCTTACGGGTCCCGCGGTGATGCGGTGCTGCTATAGCCATCGGATTCGCTTGGT	1752
QY	525	ValLeuLeuHisIleAlaIleValGlnAlaAlaLeuProGlnGlyIleValProPheVal	544
DB	1753	GATTTACTCGGTGCTGCTATAGTTCAGGCGCCATTCCTCAAGGAATGTGCCCTTTGTG	1812
QY	545	PheAlaLysGlyTyrGlyValHisProAspIleLeuSerThrAla-----TyrGly---	561
DB	1813	TTTGGAAAGGAGTCAATGTTTCATCTCTCTATTTTAAAGTACAGGGGTAAATTTTGAATG	1872
QY	562	-----ProIleThr	564
DB	1873	CTTATAGCGCTTCGATCAAG	1893
RESULT 12			
ABZ13042			
ID ABZ13042 standard; DNA; 1923 BP.			
XX			



AC ABZ13042;  
 XX  
 DT 21-JAN-2003 (first entry)  
 XX  
 DE Arabidopsis thaliana stress regulated gene SEQ ID NO 847.  
 XX  
 KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN WO200216655-A2.  
 XX  
 PD 28-FEB-2002.  
 XX  
 PF 24-AUG-2001; 2001WO-US026685.  
 XX  
 PR 24-AUG-2000; 2000US-0227866P.  
 XX  
 PR 26-JUN-2001; 2001US-0264647P.  
 XX  
 PR 22-JUN-2001; 2001US-0300111P.  
 XX  
 PA (SCRI) SCRIPPS RES INST.  
 XX  
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.  
 XX  
 PI Harper JF, Kreps J, Wang X, Zhu T;  
 XX  
 DR WPI; 2002-304127/34.  
 XX  
 PT Identifying a stress condition to which a plant cell has been exposed and  
 XX  
 PT producing plants with increased tolerance to these abiotic stresses.  
 XX  
 PS Claim 144; SEQ ID NO 847; 577pp + Sequence Listing; English.  
 XX  
 CC The invention relates to identifying a stress condition to which a plant  
 CC cell has been exposed, comprising: (a) contacting nucleic acid  
 CC representative of expressed polynucleotides in the plant cell with an  
 CC array or probes representative of the plant cell genome; and (b)  
 CC detecting a profile of expressed polynucleotides in the plant cell  
 CC characteristic of a stress response. The method is useful in the  
 CC production of transgenic plants, cells and seeds and in producing plants  
 CC with increased tolerance to abiotic stress. The present sequence is that  
 CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used  
 CC in methods of the invention. Note: The sequence data for this patent is  
 CC not represented in the printed specification but is based on sequence  
 CC information supplied to Derwent by the European Patent Office  
 XX  
 SQ Sequence 1923 BP; 459 A; 522 C; 435 G; 507 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 3,75e-100 Length: 1923  
 Score: 1560.50 Matches: 353  
 Percent Similarity: 63.87% Conservative: 73  
 Best Local Similarity: 52.92% Mismatches: 102  
 Query Match: 53.63% Indels: 139  
 DB: 6 Gaps: 20

US-10-030-884-14 (1-573) x ABZ13042 (1-1923)

QY 1 MetThrAlaLeuAspLeuThrHisValLeuThrAlaValProLeuTyrValAla 20  
 DB 1 ATGATCTCATGGACGACCTCTACACGGTCTTACCGCGGTGATTCCTCTCTACGTGGCC 60  
 QY 21 MetThrLeuAlaTyrGlySerValArgTrpTrpArgillePheThrProAspGlnCysSer 40  
 DB 61 ATGATCTCTGCTTACGGCTCTGTCCGGTGTGGAAATCTTCTCACCGGACCAATGCTCC 120  
 QY 41 GlyTleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
 DB 121 GGAATCAACCGTTTGTGCGCATCTTCGCGGTCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180  
 QY 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80  
 DB 181 ACCAACATCTTACGCAATGATCTCCGGTTCATCGCGCGGACACACTCTCCAAAAATC 240

QY 81 AlaValLeuAlaLeuLeuAlaLeuAlaSerArgGlyLeuSerSerProArgAlaLeuGly 100  
 DB 241 ATCATGTTCTCACTTTTAGTCTCTATGGGCT-----AATTTCACCTCGCTCCGGTAGC 291  
 QY 101 LeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThrLeuValMetGly 120  
 DB 292 CTCGAGTGGAGCATCAATCTTTTCCCTCTCCACACTTCCCAACACTCTTGTATATGGGG 351  
 QY 121 IleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeuMetValGlnVal 140  
 DB 352 ATTCCCTCTCTTGATCGCCCATGTATGGGAATACTCT---GGTTCCTCATGTGCCAAATC 408  
 QY 141 ValValLeuGlnCysIleIleTrpTyrThrLeuMetLeuPheLeuPheGluTyrArgAla 160  
 DB 409 GTCGTCCTCCAGTGTATCATCTGGTACAGCTTCTCTTTTCTCTCTCTCTCTCTCTCTCT 468  
 QY 161 AlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSerIleValSerPhe 180  
 DB 469 GCCAGATGCTCATCATGAGGAGTCTCCCTGAG---ACGCTGCTTCCATGTTTCTTCTTC 525  
 QY 181 ArgValAspSerAspValValSerLeuAlaArgGlyAsp---ValGluLeuGluAlaGlu 199  
 DB 526 AAAGTCGAATCCGACGTCGTTTCGCTCGACGGCCATGATTTCTTTCGAGACCCGATGAGAG 585  
 QY 200 ProAspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGlyAspAlaGlyArgVal 219  
 DB 586 ATA-----GGTGCACGCGGGAAGCTT 606  
 QY 220 ArgValThrValArgLysSerThrSerSerArgSerGluAlaAlaCysSerHisSerHis 239  
 DB 607 CACGTCAACCGTGAGAAAAATCAACGCTTCACGT---CGTTCGTTCTGC----- 651  
 QY 240 SerGlnThrMetGlnProArgValSerAsnLeuSerGlyValGluIleTyrSerLeuGln 259  
 DB 652 GGCCCGAACACTACTCCCGCGCGTCAAAATCTCACCGAGCTGAGATTTATAGTCTCAGC 711  
 QY 260 SerSerArgAsnProThrProArgGlySerSerPheAsnHisAlaAspPhePheAsnIle 279  
 DB 712 ACC-----ACTCTAGAGGCTCTAAATTCACCACTCTCATTTTACAACTG 759  
 QY 280 ValGly-----Ala 282  
 DB 760 ATGGGTTCCTCCCGGTGCTGCTCTCTCAATTTGGTCCCGCGGATATGACTCCGTTCAA 819  
 QY 283 AlaAlaLysGly-----GlyGlyGlyAlaAlaGly 292  
 DB 820 TCATCTAGAGTCCAACTCTCGACCTTCAAACTTCGAGGAGAAATTCGCGCATCGCATCC 879  
 QY 293 AspGluGluLysGlyAlaCysGlyGlyGlyGlyGlyHisSerProGlnPro----- 310  
 DB 880 TCCCGGAGATTCGGGTATTACCTCGAGGAGGAGCGGGTCTTTATCCGGCTCCGAATCCA 939  
 QY 310 ----- 310  
 DB 940 GAGTTCTTCAACCAACCACTCTACGCGCAATAAAAGCGTCAATAAAACCGGAAGAC 999  
 QY 311 -----GlnAlaValAlaValProAla-----LysArg 319  
 DB 1000 GTTAATACGAATCAGCAGACGACTCTTCCAAAGCGGCGCAAGTCAAAACAGCCATGAGCC 1059  
 QY 320 LysAspLeuHisMetLeuValTrpSerSerSerSerSerSerProValSerGluArgAlaAla 339  
 DB 1060 AAGAGCTTCACATGTCTGCTGAGCTCAACCGGTTCACCCGTTTCAGACCGGGCGGGT 1119  
 QY 340 ValHisValPheGlyAlaGlyAlaAspHis-----AlaAspValLeu 354  
 DB 1120 CTTAACGTTTTCGGC---GGAGCACCTGACACGATCAAGCGGGAAGATCTGCAAGGT 1176  
 QY 355 AlaLys----- 356  
 DB 1177 GCTAAAGAGATCGGTATGTTAGTCCAGATCAATCTCAACCGCGGAGACCAAGCTGTA 1236  
 QY 357 -----GlyAlaGlnAlaTyrAspGluTyrGlyArgAsp 367

Db	1237	GCTCATCCAGCAAGTGGAGATTTCGGAGGAGAACAACTTTAGTTTCGCCGGAAGAA	1296	PR	25-MAR-1999;	99US-0126264P.
Qy	368	AspTyrSerSerArgThrIleAsnGlySerGlyGlyAlaAspLysGlyGlyProThrLeu	387	PR	29-MAR-1999;	99US-0126785P.
Db	1297	GAGAAGCAGAGACACCAAGACCCGAGATGGT-----CTA	1335	PR	01-APR-1999;	99US-0127462P.
Qy	388	SerLysLeuGlySerAsnSerThrAlaGlnLeuTyrProLysAspAspGlyGluGlyArg	407	PR	06-APR-1999;	99US-0128234P.
Db	1336	AACAAACTTGCTCCAAATTCACCGCGCGGTACAAATCCAAG--ACAGGTCTAGGAGGA	1392	PR	08-APR-1999;	99US-0128714P.
Qy	408	AlaAlaAlaValAla-----MetProProAlaSerValMetThrArgLeuIleLeu	424	PR	16-APR-1999;	99US-0129845P.
Db	1393	GCCGAAGCAAGTCAACGAAATAATGCTCCGGGAGTGATGACAAAGCTGATATCTG	1452	PR	21-APR-1999;	99US-0130077P.
Qy	425	IleMetValTrpArgLysLeuIleArgAsnProAsnThrTyrSerSerLeuIleGlyVal	444	PR	23-APR-1999;	99US-0130449P.
Db	1453	ATAATGGTTTGAGGAAATCATCAGAACCCCAACACTTACTCTAGTCTCATTTGGACTT	1512	PR	28-APR-1999;	99US-0130891P.
Qy	445	ValTrpSerLeuValSerTyrArgTrpGlyIleGluMetProAlaIleAlaArgSer	464	PR	30-APR-1999;	99US-0131449P.
Db	1513	ATTGGGCTCTCGTCTGCTTCCGGTGGCAGCGTGGCAATGCCCAAAATCATTCACAACTCT	1572	PR	30-APR-1999;	99US-0132048P.
Qy	465	IleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeuGlyLeuPheMet	484	PR	03-MAY-1999;	99US-0132407P.
Db	1573	ATCTCATCTATCTGATGCTGGTCTTGGATGGCAATGTTAGTTGGGGTGTTCATG	1632	PR	05-MAY-1999;	99US-0132484P.
Qy	485	AlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAlaIleAlaMetGlyVal	504	PR	06-MAY-1999;	99US-0132485P.
Db	1633	GGGTTGCAACCAATTAATCCGCTTGTGGGAATTCAGTGGCAACGTTTGGCATGGCGGTT	1692	PR	06-MAY-1999;	99US-0132486P.
Qy	505	ArgPheValAlaGlyProAlaValMetAlaAlaSerIleAlaValGlyLeuArgGly	524	PR	07-MAY-1999;	99US-0132863P.
Db	1693	AGTTCCCTTACGGTCCGCGCGTGTGGCGTGTGCTGTAGCCATCGGATTACGTGGT	1752	PR	11-MAY-1999;	99US-0133425P.
Qy	525	ValLeuLeuHisIleAlaIleValGlnAlaLeuProGlnGlyIleValProPheVal	544	PR	14-MAY-1999;	99US-0134218P.
Db	1753	GATTTACTCGGTGCTGCTATAGTTACGGCGCGCATTTACCTCAAGGAATTTGCCCTTTGTG	1812	PR	14-MAY-1999;	99US-0134219P.
Qy	545	PheAlaLysGlyTyrGlyValHisProAspIleLeuSerThrAla-----TyrGly---	561	PR	14-MAY-1999;	99US-0134221P.
Db	1813	TTTGGAGAGAGTACAAATGTTCTCTCTATTTAAGTACAGGGGTAAATTTGGAAATG	1872	PR	18-MAY-1999;	99US-0134370P.
Qy	562	-----ProIleThr 564		PR	18-MAY-1999;	99US-0134768P.
Db	1873	CITATAGCGCTTCGATCATCAG 1893		PR	19-MAY-1999;	99US-0134941P.
RESULT 13				PR	20-MAY-1999;	99US-0135124P.
ID AAC43229				PR	21-MAY-1999;	99US-0135353P.
XX AAC43229;				PR	24-MAY-1999;	99US-0135629P.
XX AAC43229;				PR	25-MAY-1999;	99US-0136021P.
DT 17-OCT-2000 (first entry)				PR	28-MAY-1999;	99US-0136392P.
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 38501.				PR	01-JUN-1999;	99US-0136782P.
DE DE				PR	03-JUN-1999;	99US-0137222P.
XX Hybridisation assay; genetic mapping; gene expression control;				PR	04-JUN-1999;	99US-0137528P.
KW protein identification; signal transduction pathway; metabolic pathway;				PR	07-JUN-1999;	99US-0137724P.
KW promoter; termination sequence; ss.				PR	08-JUN-1999;	99US-0138094P.
XX Arabidopsis thaliana.				PR	10-JUN-1999;	99US-0138540P.
OS Arabidopsis thaliana.				PR	10-JUN-1999;	99US-0138847P.
XX EF1033405-A2.				PR	14-JUN-1999;	99US-0139119P.
PN PN				PR	16-JUN-1999;	99US-0139452P.
XX 06-SEP-2000.				PR	17-JUN-1999;	99US-0139453P.
XX 25-FEB-2000; 2000EP-00301439.				PR	18-JUN-1999;	99US-0139454P.
XX 25-FEB-1999; 99US-0121825P.				PR	18-JUN-1999;	99US-0139455P.
PR 05-MAR-1999; 99US-0123180P.				PR	18-JUN-1999;	99US-0139456P.
PR 09-MAR-1999; 99US-0123548P.				PR	18-JUN-1999;	99US-0139457P.
PR 23-MAR-1999; 99US-0125788P.				PR	18-JUN-1999;	99US-0139458P.

PR	16-JUL-1999;	99US-0144086P.	PR	08-OCT-1999;	99US-0158232P.
PR	19-JUL-1999;	99US-0144325P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159329P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144332P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144332P.	PR	14-OCT-1999;	99US-0159637P.
PR	20-JUL-1999;	99US-0144884P.	PR	18-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-0144814P.	PR	21-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160767P.
PR	22-JUL-1999;	99US-0145085P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145132P.	PR	22-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160989P.
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145933P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145938P.	PR	25-OCT-1999;	99US-0161406P.
PR	27-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161920P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161992P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161993P.
PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.
PR	05-AUG-1999;	99US-0147302P.	Alignment Scores:		
PR	05-AUG-1999;	99US-0147302P.	Pred. No.:	6.56e-99	Length:
PR	06-AUG-1999;	99US-0147260P.	Score:	1542.50	Matches:
PR	06-AUG-1999;	99US-0147303P.	Percent Similarity:	62.59%	Conservative:
PR	06-AUG-1999;	99US-0147303P.	Best Local Similarity:	53.53%	Mismatches:
PR	09-AUG-1999;	99US-0147935P.	Query Match:	53.01%	Indels:
PR	11-AUG-1999;	99US-0148171P.	DB:	3	Gaps:
PR	12-AUG-1999;	99US-0148319P.	US-10-030-884-14 (1-573) x AAC43229 (1-1851)		
PR	13-AUG-1999;	99US-0148565P.	QY	1	MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValValProLeuTyrValAla 20
PR	13-AUG-1999;	99US-0148684P.	Db	1	ATGATTACGTGGCAGACTTGACACCGTCCTCACCGCGTGGTACCTTTACGTAGCT 60
PR	16-AUG-1999;	99US-0149368P.	QY	21	MetThrLeuAlaTyrGlySerValArgTrrpArgIlePheThrProAspGlnCysSer 40
PR	17-AUG-1999;	99US-0149175P.	Db	61	ATGATTCTCGCCTACGGATCCGTACAGTGGTGAAGATATTCACACGACGAGTCTCC 120
PR	18-AUG-1999;	99US-0149426P.	QY	41	GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60
PR	20-AUG-1999;	99US-0149723P.	Db	121	GGCATCAACCGTTCGTGCTATCTCGCGCTCCCTCTCTCTCTCTCTCTCTCTCTCTCT 180
PR	20-AUG-1999;	99US-0149929P.	QY	61	ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnIlyVal 80
PR	23-AUG-1999;	99US-0149902P.	Db	181	ACCAACGATCCTTACGCCATGATTTCCGCTTCGTGCGCGCGACACGCTTCAAAAAATC 240
PR	23-AUG-1999;	99US-0149930P.	QY	81	AlaValLeuAlaLeuLeuAlaLeuAlaSerArgGlyLeuSerSerProArgAlaLeuGly 100
PR	25-AUG-1999;	99US-0150566P.	Db	241	ATCATGCTCGTCTTACTTGTCTATGGGCT-----AACCTAACCAAGAACGGTAGC 291
PR	26-AUG-1999;	99US-0150884P.	QY	101	LeuAspTrrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThrLeuValMetGly 120
PR	27-AUG-1999;	99US-0151065P.	Db	292	TGGAGTGGATGATCACATCTTCTCTCAGCATCTCTCCCAACACATCTTGTTCATGGGG 351
PR	27-AUG-1999;	99US-0151066P.	QY	121	IleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeuMetValGlnVal 140
PR	30-AUG-1999;	99US-0151033P.	Db	352	ATCCCTCTGTGATGCCATGTACGGA---ACCTACGACGAGTTCCTCTCTATATGGTCCAAATC 408
PR	31-AUG-1999;	99US-0151438P.	QY	141	ValValLeuGlnCysIleIleTrrpThrLeuMetLeuPheLeuPheGlnTyrArgAla 160
PR	01-SEP-1999;	99US-0151930P.	Db	409	GTGTTCCTTCAGTGTATCATTTGGTACACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 468
PR	07-SEP-1999;	99US-0153707P.			
PR	13-SEP-1999;	99US-0153758P.			
PR	15-SEP-1999;	99US-0154018P.			
PR	16-SEP-1999;	99US-0154039P.			
PR	20-SEP-1999;	99US-0154779P.			
PR	22-SEP-1999;	99US-0155139P.			
PR	23-SEP-1999;	99US-0155486P.			
PR	24-SEP-1999;	99US-0155659P.			
PR	28-SEP-1999;	99US-0156458P.			
PR	29-SEP-1999;	99US-0156598P.			
PR	04-OCT-1999;	99US-0157117P.			
PR	06-OCT-1999;	99US-0157533P.			
PR	06-OCT-1999;	99US-0157865P.			
PR	07-OCT-1999;	99US-0158029P.			

Qy	161	AlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaLaserIleValSerPhe	180
Db	469	GCTAAAGCTTCTTATTATGGACAGCTCCCGAG---ACTGGTCGGTCTATTTGTTTCGTTT	525
Qy	181	ArgValAspSerAspValValSerLeuAlaArgGlyAsp---ValGluLeuGluAlaGlu	199
Db	526	AAAGTTGATCCGACGTCGTTTCTTCGCGTCATGATTTTCTTGAGACGGATGCTGAG	585
Qy	200	ProAspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGlyAspAlaGlyArgVal	219
Db	586	ATA-----GGAAACGACGGAAAGCTT	606
Qy	220	ArgValThrValArgLysSerThrSerSerArgSerGluAlaAlaCysSerHisSerHis	239
Db	607	CATGTTACCGTAGAANAATCAACGCATCGAGA-----CGG	642
Qy	240	SerGlnThrMetGlnProArgValSerAsnLeuSerGlyValGluIleTyrSerLeuGln	259
Db	643	TCACGTGATGACTCCACGCGCTTCGAATCTTACCGGAGCTGAGATCTATAGCTTAGT	702
Qy	260	SerSerArgAsnProThrProArgLysSerSerPheAsnHisAlaAspPhePheAsnIle	279
Db	703	TCG-----ACTCCGAGAGGTTCTAACTTTAACCATTTCTGATTTTACTCTGTT	750
Qy	280	ValGlyAlaAlaAlaLysGlyGlyGlyAlaAlaGlyAsp-----	293
Db	751	ATGGGGTTTCCC-----GGCGGAGGCTTCGAATTTGGTCGGCGGATTTG	798
Qy	294	-----GluGluLysGly	297
Db	799	TACTCCGTTCAATCTTCTCGTGGTCCGACTCCACGGCTTCGAACTTCGAGAGACAAC	858
Qy	298	AlaCysGlyGlyGly-----GlyGlyGlyHisSerProGlnProGlnAlaValAlaVal	315
Db	859	GCGTTAAATATGGAATTTACAATAACACTAACAGTCTGTTCCGGCGGCGGTTCTGAC	918
Qy	316	ProAlaLysArg-----	319
Db	919	CGGCTCGAACCCCGAGTTTTCAACCGGTACCGGTGTTTCACTAAACCGAATAAAATT	978
Qy	320	-----Lys	320
Db	979	CCTAAAGAAAACCAACAGCACTCAAGAGAAAGATGCAAGCGAGCCATGACGCTAAG	1038
Qy	321	AspLeuHisMetLeuValTrpSerSerAlaSerProValSerGluArgAlaAlaVal	340
Db	1039	GAGCTTCACATGTTGTTGTTGAGCTCAAGCGCTTCTCGGTCTCCGAC-----	1086
Qy	341	HisValPheGlyAlaGlyAlaAspHisAlaAspValLeuAlaLysGlyAlaGlnAla	360
Db	1087	---GTGTTTGGCGAGGTGTCAGGCGACAACGT-----GCACCGGAA	1125
Qy	361	TyrAspGluTyrGlyArgAspAspTyr-----SerSerArgThrLys	374
Db	1126	CAATCTGAACAGGTGCTAAGAGATTCGGATGTTGTCTCTGATCAACCTCGAAGAGT	1185
Qy	375	AsnGlySerGlyAlaAsp-----	381
Db	1186	AATGCTAGAGGTGGTGGTATCATTCGCGCGTCTTGATGTGAGAGGGGAGAAAGAGAG	1245
Qy	382	-----LysGlyGlyProThrLeuSerLysLeuGlySerAsnSerThrAlaGlnLeuTyr	399
Db	1246	ATAGAGAAAGCTACAGCAGGCGCTGAATAAATGGGGTCTAATTTCCACGGCGGAGCTAGAG	1305
Qy	400	ProLysAspAspGlyGluGlyArgAlaAlaAlaValAlaMetProProAlaSerValMet	419
Db	1306	GCGGCTGTGGAGATGGCGGCGCAACCAACGAGAACACATATGCGCGCCGACAGTGTGATG	1365
Qy	420	ThrArgLeuIleLeuIleMetValTrpArgLysLeuIleArgAsnProAsnThrTyrSer	439
Db	1366	ACACACACTGATATTGATATGTGTGGAGAAAGCTGATCAGAAACCCCAACAGCTATCC	1425
Qy	440	SerLeuIleGlyValValTrpSerLeuValSerTyrArgTrpGlyIleGluMetProAla	459

[illegible]

XX DNA encoding a root-specific auxin transport protein, used to develop  
 PT transgenic plants with increased resistance to herbicides.

XX Claim 1; Fig 6; 55pp; English.

XX The present sequence is a cDNA encoding BIR1 which functions as a root-  
 CC specific auxin transport (efflux) carrier protein involved in  
 CC gravitropism. The sequence was obtained from A. thaliana cDNA phage  
 CC library. The sequence is used for producing genetically engineered plants  
 CC with greater resistance to auxin-based herbicides and auxin transport  
 CC inhibitors in combination with a second herbicide. It can also be used to  
 CC enhance transport of auxin in plant roots, produce transgenic plants  
 CC which exhibit altered auxin homeostasis and mutant plants in which the  
 CC roots are agravitropic and have reduced sensitivity to ethylene

XX Sequence 2204 BP; 592 A; 484 C; 541 G; 587 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.: 2,69e-98 Length: 2204  
 Score: 1535.00 Matches: 347  
 Percent Similarity: 60.89% Conservative: 64  
 Best Local Similarity: 51.41% Mismatches: 116  
 Query Match: 52.75% Indels: 148  
 DB: 3 Gaps: 17

US-10-030-884-14 (1-573) x AA229279 (1-2204)

QY 1 MetIleThAlaLeuAspLeuTyrHisValLeuThrAlaValProLeuTyrValAla 20  
 DB 19 ATGATCACCAGCAAGACATGATGATGTTTATGGCGCTATGTCGCCGTATACGTTGCT 78  
 QY 21 MetThrLeuAlaTyrGlySerValArgTTPAArgIlePheThrProAspGlnCysSer 40  
 DB 79 ATGATATTAGCTATGCTGCGTACGTTGGTGGGATATTACACCGGACCAATGTTTCC 138  
 QY 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
 DB 139 GGTATAAACCGGTTCGTTCGGTTCCTCGGGTTCCTCTCTCTCTCTCTCTCTCTCT 198  
 QY 61 ThrAsnAspPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnIleVal 80  
 DB 199 TCCATGATCCCTTACCAATGAATACCACTTCTCCGCTGCTGATCTCTTTCAGAAAGTC 258  
 QY 81 AlaValLeuAlaLeuLeu-----AlaLeuAlaSerArgGlyLeuSerSerPro 96  
 DB 259 GTTATCTCGCGCACTCTTTCTTTGGCAGGGGTTTAGCCGACAGGA----- 306  
 QY 97 ArgAlaLeuGlyLeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThr 116  
 DB 307 -----AGCCTAGAATGGATGATAACGCTCTTTTCACTATCAACACTGCTCTTAACAGC 357  
 QY 117 LeuValMetGlyIleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeu 136  
 DB 358 TTGGTAATGGGAATCCATTGCTTAGGGCGATGTACGGAGACTTCTCC---GGTAACCTA 414  
 QY 137 MetValGlnValValLeuGlnCysIleIleTyrThrLeuMetLeuPheLeuPhe 156  
 DB 415 ATGGTGCAGATCGTGGTGTTCAGAGCATCATATGATATATATATGCTCTCTCTCTCT 474  
 QY 157 GluTyrArgAlaAlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSer 176  
 DB 475 GAGTTCGGTGGGCTAAGCTTCTATCTCCGAGCAGTTCCTCCGAG---ACGGCTGGTTCA 531  
 QY 177 IleValSerPheArgValAspSerAspValValSerLeu---AlaArgGlyAspValGlu 195  
 DB 532 ATTACTTCTCTCAGAGTTGACTCTGATGTTATCTCTTAATGCCCGGTGAACCCCTCCAG 591  
 QY 196 LeuGluAlaGlyProAspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGlyAsp 215  
 DB 592 ACCGATGCCGAGATA-----GGAGAC 612  
 QY 216 AlaGlyArgValArgValThrValArgLysSerThrSerSerArgSerGluAlaAlaCys 235

DB 613 GACGGAAGCTACACGTGTGTGTGGAAGATCAAGTGGCGCTCATATGATCTCTTCA 672  
 QY 236 -----SerHisSer-----HisSerGlnThrMetGlnProArgValSerAsn 249  
 DB 673 TTCACCAATCTCACGGCGAGGACTTAACTCTCTCATGATATACCGCGAGCTTCAAT 732  
 QY 250 LeuSerGlyValGluIleTyrSerLeuGlnSerSerArgAsnProThrProArgGlySer 269  
 DB 733 CTCACCGGGGTAGATTTTACTTCGTTCAATCGTCACGAGAGCCGACCGAGAGCTTCT 792  
 QY 270 SerPheAsnHisAlaAspPhePhe----- 277  
 DB 793 AGCTTTAATCAGACAGATTTCTACGCAATGTTTAACGCAAGCAAAAGCTCCAAAGCCTCGT 852  
 QY 278 -----AsnIleValGlyAlaAlaLysGlyGly----- 288  
 DB 853 CACGGTTACACTAATAGCTACGGCGCGCTGAGCTGCTCCAGGTGGAGATGTTTACTCA 912  
 QY 288 ----- 288  
 DB 913 CTTCACTCTTCTAAAGCGGTGACGCGGAGAACCTCAATTTTGTATGAGGAAGTTATGAAG 972  
 QY 288 ----- 288  
 DB 973 ACGCGCAAGAAAGCAGAGAGGAGGAGATGATGAGTGGGGAATTATACACCAATAT 1032  
 QY 289 -----GlyAlaAla 291  
 DB 1033 AGTGTTCGTGCTACCCACCGCGGACCAATGTTACGGGGTCAACGAGTGGAGCAAGT 1092  
 QY 292 GlyAspGluGlyGlyAlaCysGlyGlyGlyGlyGlyGlyHisSerProGlnProGln 311  
 DB 1093 GGAGTCAAGAAAGAAAGTGTGTGGGAGGAGCGGTGCG----- 1134  
 QY 312 AlaValAlaValProAlaLysArgLysAspLeuHisMetLeuValTrpSerSerSerAla 331  
 DB 1135 CGAGTAGGAGTAGAGGAGCAAAACAGAGAGATGAACATGTTCTGTGGAGTTCAGTGTCT 1194  
 QY 332 SerProValSerGluArgAlaAlaValHisValPheGlyAlaGlyGlyAla----- 348  
 DB 1195 TCTCGGTGTGGAAGCCACCGCAAGAAATGCTATGACAGAGGTCTCTCCACCGATGTA 1254  
 QY 349 -----AspHisAlaAspValLeuAlaLysGlyAlaGln 359  
 DB 1255 TCCACCGACCTTAAGTTTCTATTCTCTCTCACGCAACCTGGCTACTAAAGCGATGAG 1314  
 QY 360 AlaTyrAspGluTyr-----GlyArgAspAspTyrSerSerArgThrLysAsnGly 376  
 DB 1315 AATCTGTATAGAAACATGTCTACCGGAGAAAGAGGCGATGTGGAATGGACACAGCGGT 1374  
 QY 377 SerGlyGlyAlaAspLysGlyGlyProThrLeuSerLysLeuGlySerAsnSerThrAla 396  
 DB 1375 ANTAACGGGGGAAG-----TCACCTTACATGGGCAAAAGGTAGGAC----- 1419  
 QY 397 GlnLeuTyrProLysAspGlyGluGlyArgAlaAlaValAlaMetProProAla 416  
 DB 1420 -----GTGGAAGACGGCGGTCCCGGTCTTAGGAAACAGCAGATGCCCGCGCG 1467  
 QY 417 SerValMetThrArgLeuIleLeuIleMetValTrpArgLysLeuIleArgAsnProAsn 436  
 DB 1468 AGTGATGATGAGAGACTAATTTCTGATATGTTTGGAGAAACTCATTCGAAACCTTAAC 1527  
 QY 437 ThrTyrSerSerLeuIleGlyValValTrpSerLeuValSerTyrArgTTPGlyIleGlu 456  
 DB 1528 ACTTACTCTAGTCTTTGGCCTTGTGGTCTTGTGGTCTTGTCTCTCAAGTGGATATAAAG 1587  
 QY 457 MetProAlaIleIleAlaArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAla 476  
 DB 1588 ATGCCCAACGATAATGAGTGGATCGATTTCGATATTAATCTGATGCTGTCTTGGATGCT 1647  
 QY 477 MetPheSerLeuGlyLeuPheMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLys 496

Db 1648 ATGTTTAGCTTGGTCTATTTATGCAATTCGCAACCAAGATTATTGGTGGGAAATCA 1707  
 Qy 497 LeuAlaAlaIleAlaMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaAla 516  
 Db 1708 GTAGCAGGGTTGGCATGCCGTAGGTTCTTGACTGACACGCGGTGATCGCAGCCACC 1767  
 Qy 517 SerIleAlaValGlyLeuArgGlyValLeuLeuHisIleAlaIleValGlnAlaAlaLeu 536  
 Db 1768 TCAATAGCAATGGTATTCGAGGTGATCTCTCCATATCGCATCGTTCAGGCTGCTCT 1827  
 Qy 537 ProGlnGlyIleValProPheValPheAlaGlyGlyValHisProAspIleLeu 556  
 Db 1828 CTTCAAGAGATCGTTCTCTTTTGGTTTCCCAAGATATTAACGTCATCTGATATCTC 1887  
 Qy 557 SerThrAla-----TyrGly-----ProIleThr 564  
 Db 1888 AGCACTGGCGGTATATTCGGAATGCTGGTCTTGTGCTGTGAACA 1932

## RESULT 15

AAZ57348

ID AAZ57348 standard; cDNA to mRNA; 2232 BP.

XX AAZ57348;

XX DT 03-APR-2000 (first entry)

XX DE Gravitropism stimulation response related protein AGR cDNA sequence.

XX KW Gravitropism stimulation response; AGR; plant root; ds.

XX OS Arabidopsis thaliana.

XX FH Key Location/Qualifiers

XX FT CDS 24..1967

XX FT /\*tag= a

XX FT /product= "AGR"

XX FT /note= "a gravitropism stimulation response related

XX FT protein"

XX JPI1318463-A.

XX PN 24-NOV-1999.

XX PD 15-MAY-1998; 98JP-00134097.

XX PF 15-MAY-1998; 98JP-00134097.

XX PR (KAZU-) ZH KAZUSA DNA KENKYUSHO.

XX PA (OJIP) OJI PAPER CO.

XX PA WPI; 2000-100768/09.

XX DR P-PSDB; AAY53129.

XX XX

XX PT A gene participating to the gravitropism stimulation response of a plant

XX PS root.

XX PS Claim 2; Page 11-14; 20pp; Japanese.

XX CC The present sequence encodes a protein which participates in the

XX CC gravitropism stimulation response of a plant root. The protein is

XX CC designated AGR. AGR can be used to improve the fixing rate of a plant

XX CC root to soil

XX SQ Sequence 2232 BP; 601 A; 484 C; 544 G; 603 T; 0 U; 0 Other;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

US-10-030-884-14 (1-573) x AAZ57348 (1-2232)

Qy 1 MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValValProLeuTyrValAla 20  
 Db 24 ATGATCCACCGGCAAGACATGTACGATGTTTACGGCTATGGTGGCTATACGTTGCT 83  
 Qy 21 MetThrLeuAlaTyrGlySerValArgTyrTrpArgIlePheThrProAspGlnCysSer 40  
 Db 84 ATGATATTAGCTATGTTCCGTACGTTGGTGGGATATTCACACCGGACCAATGTTCC 143  
 Qy 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
 Db 144 GGTATAACCGGTTCGTTGGGTTTTCGGGTTTCCTCTCTCTCTCTCTCTCTCTCTCTCT 203  
 Qy 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnIleVal 80  
 Db 204 TCCAAATGATCTTATGCAATGAATTACCACTTCTCTGCTGCTGATCTCTTCAGAAATC 263  
 Qy 81 AlaValLeuAlaLeuLeu-----AlaLeuAlaSerArgGlyLeuSerSerPro 96  
 Db 264 GTTATCTCCCGCCTCTCTTTCTTGGCAGGCGTTTTCGCGGAGAGG----- 311  
 Qy 97 ArgAlaLeuGlyLeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThr 116  
 Db 312 -----AGCCTAGAAATGATGATAACGCTCTTTTCACTATCAACACTGCTTAACAG 362  
 Qy 117 LeuValMetGlyIleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeu 136  
 Db 363 TTGTAATGGGAATCCCATTTAGGGCATGTACGGAGACTTCTCC---GGTAACCTA 419  
 Qy 137 MetValGlnValValLeuGlnCysIleIleTyrThrLeuMetIlePheLeuPhe 156  
 Db 420 ATGTGTCAGATCGTGGTGTTCAGAGCATCATATGATATACATTAAATGCTCTCTTCTTT 479  
 Qy 157 GluTyrArgAlaAlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSer 176  
 Db 480 GAGTTCGTCGGGCTAAGCTTCTCATCTCCGAGCAGTCTCCCGAG---ACGCTGTTCA 536  
 Qy 177 IleValSerPheArgValAspSerAspValValSerLeu---AlaArgGlyAspValGlu 195  
 Db 537 ATTACTTCTTCAGAGTTGACTCTGATGTATCTCTCTTAATGGCGGTGAACCTCCAG 596  
 Qy 196 LeuGluAlaGluProAspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGlyAsp 215  
 Db 597 ACCGATCGGAGATA-----GGAGAC 617  
 Qy 216 AlaGlyArgValArgValThrValArgLysSerThrSerSerArgSerGluAlaAlaCys 235  
 Db 618 GACGAAAGGTACACGCTGGTGTTCGAAGATCAAGTCCGCTCATCATGATCTCTTCA 677  
 Qy 236 -----SerHisSer-----HisSerGlnThrMetGlnProArgValSerAsn 249  
 Db 678 TTCACAAATCTCACGGCGAGGACTTAACTCTCTCATGATTAACCGCGAGCTTCAAT 737  
 Qy 250 LeuSerGlyValGluIleTyrSerLeuGlnSerSerArgAsnProThrProArgGlySer 269  
 Db 738 CTCACGGCGGTGAGATTTACTCCGTTCAATCGTCACGAGAGCCGACGCGGAGAGCTTCT 797  
 Qy 270 SerPheAsnHisAlaAspPhePhe----- 277  
 Db 798 AGTTTAAATCAGACAGATTTCTACGCAATGTTTAAACGCAAGCAAGCTCCAAGCCTCGT 857  
 Qy 278 -----AsnIleValGlyAlaAlaLysGlyGly----- 288  
 Db 858 CACGGTTTACACTAATAGTACGCGCGCTGGAGTGGTCCAGTGGAGATGTTTACTCA 917  
 Qy 288 ----- 288  
 Db 918 CTTCACTCTTCTTAAGCGCGTACGCGGAGAACGTCAAATTTTGTATGAGAGTTATGAAG 977  
 Qy 288 ----- 288  
 Db 978 ACGCGGAAGAAAGCAGGAAGAGGAGCAGAGATGATGAGTGGGGAATATATCAACAATAAT 1037

## Alignment Scores:

Pred. No.:	2.73e-98	Length:	2232
Score:	1535.00	Matches:	347
Percent Similarity:	60.89%	Conservative:	64
Best Local Similarity:	51.41%	Mismatches:	116
Query Match:	52.75%	Indels:	148
Ds:	3	Gaps:	17

QY 289 -----GlyAlaAla 291  
Db 1038 AGTGTTCGTCGTACCCACCGCCGAAACCCAAATGTTCAACGGGTCAACAGAGTGGAGCAAGT 1097  
QY 292 GlyAspGluGlyGlyAlaCysGlyGlyGlyGlyGlyHisSerProGlnProGln 311  
Db 1098 GGAGTCAAGAAAGGAAAGTGTGGCGGAGGAGCGGTGGC----- 1139  
QY 312 AlaValAlaValProAlaLysArgLysAspLeuHisMetLeuValTrpSerSerAla 331  
Db 1140 GGAGTAGGAGTAGGAGGACAAACAGAGAGATGAACATGTTCTGTTGGAGTTCGAGTGCT 1199  
QY 332 SerProValSerGluArgAlaAlaValHisValPheGlyAlaGlyGlyAla 348  
Db 1200 TCTCCGGTGTGGGAAGCCACGCGAAGAAATGCTATGACAGAGGTTCTTCCACCGATGTA 1259  
QY 349 -----AspHisAlaAspValLeuAlaLysGlyAlaGln 359  
Db 1260 TCCACCGACCTAAAGTTTCTATTCTCTCACGACAACCTCGCTACTAAAGCGATGCAG 1319  
QY 360 AlaTyrAspGluTyr-----GlyArgAspAspTyrSerSerArgThrLysAsnGly 376  
Db 1320 AATCTGATAGAACATGTCACCGGAGGAAAGGCGATGTGGAAATGGACCAAGACGGT 1379  
QY 377 SerGlyGlyAlaAspLysGlyGlyProThrLeuSerLysLeuGlySerAsnSerThrAla 396  
Db 1380 AATAACGGGGGAAG-----TCACCTTACATGGGCAAAAAGGTACGCAC----- 1424  
QY 397 GlnLeuTyrProLysAspAspGlyGluGlyArgAlaAlaValAlaMetProAla 416  
Db 1425 -----GTGGAAGACGGCGGTCCCGTCTCTAGGAACACAGCAGATGCCCGCGCG 1472  
QY 417 SerValMetThrArgLeuLeuLeuMetValTrpArgLysLeuLeuLeuArgAsnProAsn 436  
Db 1473 AGTGTGATGACGAGACTAATTTCTGATATGTTTGGAGAAACCTCATTCGAAACCCCTAAC 1532  
QY 437 ThrTyrSerSerLeuLeuGlyValValTrpSerLeuValSerTyrArgTrpGlyIleGlu 456  
Db 1533 ACTTACTCTAGTCCTTTGGCGCTTGCTTGTCCTTGCTCTCTTCAAGTGGAAATAAAG 1592  
QY 457 MetProAlaIleAlaArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAla 476  
Db 1593 ATGCCACGATAAATGAGTGGATCGATTTTCGATATTTCTGATGCTGGTCTTGGAAATGGCT 1652  
QY 477 MetPheSerLeuGlyLeuPheMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLys 496  
Db 1653 ARGTTTAGTCTTGGTCTATTATGCGATTGCAACCAAGATTATGCGTGGGAAATCA 1712  
QY 497 LeuAlaAlaIleAlaMetGlyValArgPheValAlaGlyProAlaValMetAlaAla 516  
Db 1713 GTAGCAGGGTTGGCATGCGGTAAAGTTCTTGACTGGACCGCGTGTATCCGAGCCACC 1772  
QY 517 SerIleAlaValGlyLeuArgGlyValLeuLeuHisIleAlaIleValGlnAlaAlaLeu 536  
Db 1773 TCAATAGCAATTTGGTATTGAGGTGATCTCTCCATATCGGCATCGTTCAGGCTGCTCTT 1832  
QY 537 ProGlnGlyIleValProPheValPheAlaLysGluTyrGlyValHisProAspIleLeu 556  
Db 1833 CTTCAAGGAATCGTTCTTCTTTTTCGCCAAAGAAATATAACGTCATCTCTGATATCTC 1892  
QY 557 SerThrAla-----TyrGly-----ProIleThr 564  
Db 1893 AGCACTGCGGTATATTCGAATGCTGGTGTGTTTGGCTGTGAACA 1937



OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: March 3, 2004, 09:09:54 ; Search time 6280 Seconds  
(without alignments)  
3954.707 Million cell updates/sec  
Title: US-10-030-884-14  
Perfect score: 2910  
Sequence: 1 MITALDIYHVLTVAVPLVYA.....DILSTAYGPITSHGFITCHS 573

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0  
Searched: 3470272 seqs, 21671516995 residues  
Total number of hits satisfying chosen parameters: 6940544  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US10030884/runat\_24022004\_103922\_1948/app\_query.fasta\_1.711  
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=spct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10030884 @cgn 1 1 2586 @runat\_24022004\_103922\_1948 -NCPU=3  
-NO.MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vl.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*

29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pln.\*  
35: em.htg.fod.\*  
36: em.htg.mam.\*  
37: em.htg.vrt.\*  
38: em.sy.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Match	Length	ID	Description
1	2910	100.0	2769	6	AX046826	Sequence
2	2137.5	73.5	2233	8	AK103181	Oryza sat
3	2129.5	73.2	1926	6	AX653844	Sequence
4	2020.5	69.4	1686	6	AX652923	Sequence
5	1977	67.9	91883	8	CNS080CDY	Oryza sat
6	1959.5	67.3	141025	2	AC123528	Oryza sat
7	1785.5	61.4	2297	8	AF515435	Populus t
8	1685	57.2	2358	8	AF056027	Oryza sat
9	1665	57.2	2402	8	AK102343	Oryza sat
10	1659.5	57.0	2457	8	AK101504	Oryza sat
11	1656	56.9	2343	8	AK099634	Oryza sat
12	1655.5	56.9	2126	8	AK063976	Oryza sat
13	1655.5	56.9	2470	8	AK103208	Oryza sat
14	1647.5	56.6	1827	6	AK654121	Sequence
15	1647.5	56.6	1869	8	AY093960	Arabidops
16	1647.5	56.6	2262	8	AF372950	Arabidops
17	1647	56.6	2262	8	AB085897	Cucumis s
18	1643	56.5	2549	6	AX046842	Sequence
19	1641.5	56.4	2276	8	AF089084	Arabidops
20	1641.5	56.4	2292	6	A68597	Sequence 11
21	1641.5	56.4	2292	6	A79382	Sequence 11
22	1623.5	55.8	9561	2	AC135795	Medicago
23	1614.5	55.5	2473	8	AF190881	Populus t
24	1605.5	55.2	2790	8	AF515434	Populus t
25	1599	54.9	1952	8	AY302060	Populus t
26	1596	54.8	1845	6	AX652849	Sequence
27	1592	54.7	1884	6	AX653686	Sequence
28	1587	54.5	1800	8	AY222857	Pisum sat
29	1571.5	54.0	2276	8	AK101191	Oryza sat
30	1569	53.9	2522	8	AB112364	Pisum sat
31	1566.5	53.8	2293	6	AX046850	Sequence
32	1566.5	53.8	2293	8	BT008949	Triticum
33	1560.5	53.6	1923	6	AX506152	Sequence
34	1560.5	53.6	2380	8	BT002085	Arabidops
35	1560.5	53.6	2489	8	AY136327	Arabidops
36	1560.5	53.6	2499	8	AF087818	Arabidops
37	1558.5	53.6	2462	8	BT0249297	Brassica
38	1542.5	53.0	2358	8	BT008691	Arabidops
39	1542.5	53.0	2404	8	AY091009	Arabidops
40	1541.5	53.0	2364	8	AF087016	Arabidops
41	1539.5	52.9	91870	8	AP004800	Oryza sat
42	1539	52.9	2316	8	BT0249298	Brassica
43	1535	52.7	1944	8	BT001152	Arabidops
44	1535	52.7	2196	8	AF056026	Arabidops
45	1535	52.7	2232	6	E31424	Gene partic

ALIGNMENTS

RESULT 1

AX046826  
LOCUS AX046826 2769 bp DNA linear PAT 15-DEC-2000  
DEFINITION Sequence 13 from Patent WO0068389.  
ACCESSION AX046826  
VERSION AX046826.1 GI:11876320  
KEYWORDS Zea mays  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
Clade; Panicoideae; Andropogoneae; Zea.  
1  
OROSCO, B.M., MENG, Z., BRUCE, W.B., CAHOON, R.E. and TAO, Y.  
AUXIN TRANSPORT PROTEINS  
Patent: WO 0068389-A 13 16-NOV-2000;  
E.I. DU PONT DE NEMOURS AND COMPANY (US); Pioneer Hi-Bred  
International, Inc. (US)  
FEATURES  
source  
Location/Qualifiers  
1..2769  
/organism="Zea mays"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:4577"  
ORIGIN  
Alignment Scores:  
Pred. No.: 4.5e-155 Length: 2769  
Score: 2910.00 Matches: 573  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-10-030-884-14 (1-573) x AX046826 (1-2769)  
QY 1 MetIleThAlaLeuAspLeuTyrHisValLeuThrAlaValValProLeuTyrValAla 20  
DB 171 ATGATCACCGCGTGGACCTCTACACGCTGCTGACGCGGTGTGCGGTGTACGTGGCC 230  
QY 21 MetThrLeuAlaTyrGlySerValArgTyrTrpArgIlePheThrProAspGlnCysSer 40  
DB 231 ATGACGCTGGCGTACGGCTCCGTCCTGTGGCGCATCTTACGCGGACCAAGTGTCTCC 290  
QY 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
DB 291 GGGATCAACCGCTTCGTGGCGCTTTCGCGCTGCGCTCCTCTCTTCCATCTCATCTCC 350  
QY 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnIleVal 80  
DB 351 ACCAAGACCCCTTCGCCATGAACCTGCGCTTCTTGGCCCGGACACGCTGCAGAGGTG 410  
QY 81 AlaValLeuAlaLeuLeuAlaLeuAlaSerArgGlyLeuSerSerProArgAlaLeuGly 100  
DB 411 GCGTCTCTCGCGTGTCTGGCGCTTGGCTCCCGCGGCTCTCTCTCCCGCGCGCTCGGG 470  
QY 101 LeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThrLeuValMetGly 120  
DB 471 CTGACTGGAGCATCACGCTCTTCTCCCTTCCACGCTCCCAACGCTGTCATGGGC 530  
QY 121 IleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeuMetValGlnVal 140  
DB 531 ATCCCGTCTGCGAGCAATGACGGCGCTGCTCGCGCGGACGCTCATGGTCCAGGTC 590  
QY 141 ValValLeuGlnCysIleIleTrpTyrThrLeuMetLeuPheLeuPheGluTyrArgAla 160  
DB 591 GTGCTCTCCAGTGATCATCTGGTACACGCTATGCTTCTCTTCTTCAGTACCGCGCC 650  
QY 161 AlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSerIleValSerPhe 180  
DB 651 GCGCGCGCTGCTGCTGACCAAGTTCCTCCGAGCGGCGCCCGGCTCCATGCTCTCTTC 710  
QY 181 ArgValAspSerAspValValSerLeuAlaArgGlyAspValGluLeuGluAlaGluPro 200  
DB 711 CGGCTCGACTCCGACGCTGCTCTGCTCGCCAGGGGACGCTCGAGCTCGAGGCGGAGCCC 770

201 AspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGlyAspAlaGlyArgValArg 220  
DB 771 GACGGCGTCCCGCGCGCGCGCGCTCTCTCTCCCGCGCGCGCGCGCGCGCGCGCGCG 830  
QY 221 ValThrValArgGlySerThrSerArgSerGluAlaAlaCysSerHisSerHisSer 240  
DB 831 GTACCGCTGGCAGTCCACAGCTCGCTCCGAGCGCGCTGCTCGACTCGACTCC 890  
QY 241 GlnThrMetGlnProArgValSerAsnLeuSerGlyValGluIleTyrSerLeuGlnSer 260  
DB 891 CAGACCATGCGAGCCCGCTGTCCACCTCTCCGCGGTGGAGATCTACTCGCTCGAGTCG 950  
QY 261 SerArgAsnProThrProArgGlySerSerPheAsnHisAlaAspPheAsnIleVal 280  
DB 951 TCGCGCAACCCACCCCGCGCGGTCCAGCTTCAACACCCCGACTTCTTCAACATCGTC 1010  
QY 281 GlyAlaAlaAlaGlyGlyGlyGlyGlyAlaAlaGlyAspGluGluGlyGlyAlaCysGly 300  
DB 1011 GCG 1070  
QY 301 GlyGlyGlyGlyGlyHisSerProGlnProGlnAlaValAlaValProAlaIleAsnGly 320  
DB 1071 GCGCGCGCGCGAGGACATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1130  
QY 321 AspLeuHisMetLeuValTrpSerSerAlaSerProValSerGluArgAlaAlaVal 340  
DB 1131 GACCTGCACATGCTGCTGAGCTCCAGCGCTCGCGCGCTCGCGCGCGCGCGCGCGCG 1190  
QY 341 HisValPheGlyAlaGlyAlaGlyAlaAspHisAlaAspValLeuAlaGlyAlaGlnAla 360  
DB 1191 CACGCTCTCGCGCGCGCGCGCGTGCATGCGCGCGCTCTCGCGCGCGCGCGCGCGCG 1250  
QY 361 TyrAspGluTyrGlyArgAspAspTyrSerSerArgThrIleValAsnGlySerGlyValAla 380  
DB 1251 TACACGAGTACCGCGCGCGCGCGCTACGACGACGACGACGACGACGACGACGACGAC 1310  
QY 381 AspIleGlyGlyProThrLeuSerLysLeuGlySerAsnSerThrAlaGlnLeuTyrPro 400  
DB 1311 GACAAGCGCGCGCGCGCGCTGTGCAAGCTGGGCTTCAACTCGACGCGCGCGCTGTACCCC 1370  
QY 401 LysAspAspGlyGluGlyArgAlaAlaAlaValAlaMetProProAlaSerValMetThr 420  
DB 1371 AAGACACGCGCGAGGCG 1430  
QY 421 ArgLeuIleLeuIleMetValTrpArgLysLeuIleArgAsnProAsnThrTyrSerSer 440  
DB 1431 CGGCTCATCTCTCATGTTGTGGAGAGCTGATCCGGAACCCCAACACCTACTCCAGC 1490  
QY 441 LeuIleGlyValValTrpSerLeuValSerTyrArgTrpGlyIleGluMetProAlaIle 460  
DB 1491 CTCATCGCGCTGCTGCTGCTCTCTTACAGTGGGCGCATCGAGATCGCGCGCGATC 1550  
QY 461 IleAlaArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeu 480  
DB 1551 ATCCCGCGTGGATTCGATCCTGTTCGAGCGGGGTCTCGGGATGGCCATGTTTCAGCCTA 1610  
QY 481 GlyLeuPheMetAlaLeuGlnProArgIleIleAlaCysGlyAsnIleLeuAlaAlaIle 500  
DB 1611 GGCCTGTTTCATGGCGCTGCAGCGCGAGGATCATCGCGTGGGGAACAAGTGGCGCGCATC 1670  
QY 501 AlaMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaAlaSerIleAlaVal 520  
DB 1671 CGATGGCGCTCCGCTGCTGCGAGCGCGCGCGGTCTGCGCGCGCGCTCCATCGCGCTC 1730  
QY 521 GlyLeuArgGlyValLeuLeuHisIleAlaIleValGlnAlaAlaLeuProGlnGlyIle 540  
DB 1731 GGTCTGCGCGCGTCTCTCTCCATCGCCATCGCTCCAGGCTGCTCTGCTCTCAGGGGATC 1790  
QY 541 ValProPheValPheAlaIleGluTyrGlyValHisProAspIleLeuSerThrAlaTyr 560  
DB 1791 GTGCGCTTGTGTTTCGCGCAGAGTACGGCTTCTATCCCGACATCTCTGAGCAGCGCTAT 1850

Qy 561 GlyProfilThrSerHisGlyPheIleThrCysHisSer 573  
 |||||  
 Db 1851 GGTCCATAACATCGCATGGTTTCATCAGTTCGCATAGT 1889

RESULT 2  
 AKI03181 2233 bp mRNA linear PLN 24-JUL-2003  
 LOCUS Oriza sativa (japonica cultivar-group) cDNA clone:J033121124, full  
 DEFINITION insert sequence.  
 ACCESSION AKI03181  
 VERSION AKI03181.1 GI:32988390  
 KEYWORDS FULL CDNA: CAP trapper.  
 SOURCE Oriza sativa (japonica cultivar-group)  
 ORGANISM Oriza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzeae; Oryza.

1  
 REFERENCE  
 AUTHORS  
 The Rice Full-Length cDNA Consortium, National Institute of  
 Agrobiological Sciences Rice Full-length cDNA Project Team,  
 Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,  
 Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,  
 Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,  
 Ohtsuki, K., Shishiki, T., Foundation of Advancement of International  
 Science Genome Sequencing & Analysis Group: Ootomo, Y., Murakami, K.,  
 Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,  
 Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,  
 Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,  
 Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,  
 Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN;  
 Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,  
 Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,  
 Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y.,  
 Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
 Yoshino, M., and Hayashizaki, Y.  
 Collection, mapping, and annotation of over 28,000 cDNA clones from  
 japonica rice  
 Science 301 (5631), 376-379 (2003)  
 22752273  
 12869764

2 (bases 1 to 2233)  
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,  
 Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,  
 Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoaka, T.,  
 Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,  
 Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,  
 Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,  
 Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,  
 Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
 Koya, S., Kurihara, C., Kuroseki, T., Kusumegi, T., Li, C., Lu, M.,  
 Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,  
 Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,  
 Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K.,  
 Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,  
 Osato, N., Ota, Y., Ootomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,  
 Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,  
 Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,  
 Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,  
 Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,  
 Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,  
 Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and  
 Yoshimura, A.  
 Direct Submission  
 Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of  
 Agrobiological Sciences, Department of Molecular Genetics, Head of  
 Laboratory of Gene Expression; 2-1-2 Kannondai, Teukuba, Ibaraki  
 305-8602, Japan [E-mail: skikuchi@nias.affrc.go.jp,  
 Tel:81-29-838-7007, Fax:81-29-838-7007]  
 This clone is one of the 28K full-length cDNA clones from japonica  
 rice.  
 URL : <http://cdna01.dna.affrc.go.jp/cDNA/>  
 NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K.,  
 Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,

TITLE  
 JOURNAL

COMMENT

Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,  
 Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., and  
 Yamamoto, M.  
 FAIS Genome Sequencing & Analysis Group: Ootomo, Y., Iida, Y.,  
 Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,  
 Kodama, T., Kuroseki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,  
 Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S.,  
 Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,  
 Yoshimura, A., Matsubara, K. and Murakami, K.  
 and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,  
 Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,  
 Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,  
 Hiroaka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,  
 Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,  
 Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
 Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N.,  
 Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,  
 Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,  
 Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,  
 Yasunishi, A. and Hayashizaki, Y.

# FEATURES

Location/Qualifiers

1..2233  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /clone="J033121124"

## ORIGIN

Alignment Scores:  
 Pred. No.: 1..19e-111 Length: 2233  
 Score: 2137.50 Matches: 457  
 Percent Similarity: 81.86% Conservative: 26  
 Best Local Similarity: 77.46% Mismatches: 36  
 Query Match: 73.45% Indels: 71  
 DB: 8 Gaps: 17

US-10-030-884-14 (1-573) x AKI03181 (1-2233)

Qy 1 MetIleThrAlaLeuAspLeuThrHisValLeuThrAlaValProLeuTyrValAla 20  
 |||||  
 Db 222 ATGATCAGCGTGGTGAGCTTACACACGCTTCACGCGGCTGGTGGCTGTACGTGGG 281  
 |||||  
 Qy 21 MetThrLeuAlaTyrGlySerValArgTrpTrpArgPheThrProAspGlnCysSer 40  
 |||||  
 Db 282 ATGACGCTGGCGTACGCTCGTGGTGGTGGCGCATCTTCTCCCGACGACGCTCC 341  
 |||||  
 Qy 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
 |||||  
 Db 342 GGCATCAACCGCTTCGTCGCGCTTTCGCGCGTCCGCTCTCTCTCCACTTCATCTCC 401  
 |||||  
 Qy 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnIysVal 80  
 |||||  
 Db 402 ACCACACACCGCTTCGCCATGACCTCCGCTTCCTCGCGCGCGACACGCTCCAGAAGCTC 461  
 |||||  
 Qy 81 AlaValLeuAlaLeuLeuAla-----LeuAlaSerArgGlyLeuSerSerPro 96  
 |||||  
 Db 462 ATCGTCTCGCGCTCTCGCCCTCTGTCGCGCGCTCTCGCGCGCGCGG----- 509  
 |||||  
 Qy 97 ArgAlaLeuGlyLeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThr 116  
 |||||  
 Db 510 -----TCATCGACTGGCTGCATCACCGCTTCTCTCCCTCTCCACCTCCCAACACC 560  
 |||||  
 Qy 117 LeuValMetGlyIleProLeuLeuArgGlyMetTyrGlyAlaSerSer-----Ala 133  
 |||||  
 Db 561 CTCGTATGGGACATCCGCTGTCTCAAGGGGATGTACGCGCGCGCGCGCGCGCTGACTCC 620  
 |||||  
 Qy 134 GlyThrLeuMetValGlnValValLeuGlnCysIleIleIleTrpTyrThrLeuMetLeu 153  
 |||||  
 Db 621 GGCAGCGCTCATGTGTCAGATCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 680  
 |||||



Db 289 -----TCWCTCGAGTGGCTCATCACCCCTCTTCTCCCTCTCCACCCCTCCCAACACC 339  
 Qy 117 LeuValMetGlyLeuProLeuLeuArgGlyMetTyrGlyAlaSerAla----- 133  
 Db 340 CTCGTATGGCATCCCGCTCTCAAGGGATGTACCGCCGCGCGCGCGCGCGCGCG 399  
 Qy 134 -----GlyThrLeuMetValGlnValValValLeuGlnCysIleIleTyrThr 150  
 Db 400 GCGACTCCGCGCGCTCATGGTCAGATCGTCTGTCTCCAGTGCATCATCTGGTACAG 459  
 Qy 151 LeuMetLeuPheLeuPheGlyTyrArgAlaAlaAlaAlaLeuValLeuAspGlnPhePro 170  
 Db 460 CTCATGCTGTCTCTTCAGTACCGYGCGCGCGGCTTCTGTGTCATGAGCAGTTCGCC 519  
 Qy 171 AspGlyAlaAlaAlaSerIleValSerPheArgValAspSerValValSerLeuAla 190  
 Db 520 GAC---ACCGCGCGCTCATCGTCTCTTCGGGTGGATCCGACGTCGCTCGCTGCC 576  
 Qy 191 ArgGlyAsp-----ValGluLeuGluAlaGluProAspGlyValAlaGlyAlaGly 207  
 Db 577 GCGCGCGTGTGGAGCGCGGAGCTGCGAGCGGAG----- 612  
 Qy 208 AlaValSerSerArgGlyGlyAspAlaGlyArgValArgValThrValArgLysSerThr 227  
 Db 613 -----GCGAGGTCCGGGCGATGCGAGATGCGGGTCACTGCGCGCAGTCCGACG 663  
 Qy 228 SerSerArgSerGluAlaAlaCysSerHis-----SerHisSerGlnThrMetGln 244  
 Db 664 AGCTCGCGCTCGAGCGCGGCTGCTCGCACGGGCGCGAGTCACTCGAGTCCATGCG 723  
 Qy 245 ProArgValSerAsnLeuSerGlyValGluLeuTyrSerLeuGlnSerSerArgAsnPro 264  
 Db 724 CCGCGGTCTCAACCTCTCGCGCTCGAGATTTACTCGTGCAGTGTCTCGCGAAGCCG 783  
 Qy 265 ThrProArgGlySerSerPheAsnHisAlaAspPheAsnIleValGlyAlaAlaAla 284  
 Db 784 AGCGCGGTGGCTCCAGCTTCAACCGCGGAGTCTTCAACATCGTC----- 831  
 Qy 285 LysGlyGlyGlyAlaAlaGlyAspGluLysGlyAlaCysGlyGlyGlyGly 304  
 Db 832 -----GGCAACGGCAAGCAGCGCGCAGGAGGAGGAGGCGCC-----GCGGTGGCGGG 879  
 Qy 305 GlyHisSerProGlnProGlnAlaValAlaValProAlaLysArgLysAspLeuHisMet 324  
 Db 880 GGCCACTCGCCGACCGC-----GTGGTGGGAGAGAGGAGGACCTGCACATG 927  
 Qy 325 LeuValTrpSerSerAlaSerProValSerGluArgAla----- 338  
 Db 928 TTCGTGTGAGTCAAGCGCTCGCGGTTCGGAGCGCGCGCGCAGCAGCTGGC 987  
 Qy 339 AlaValHisValPheGlyAlaGlyAlaAspHisAlaAspValLeuAlaLysGlyAla 358  
 Db 988 GCGGTGCAGCTTCGCGCGGTGTGCGCGCACCGACCGCGAC-----GCCAAGGAGCT 1041  
 Qy 359 GlnAlaTyrAspGluTyrGlyArgAspTyrSerSerArgThrLysAsnGlySerGly 378  
 Db 1042 CAGGCTATGATGAG-----TACAGTTCCGGGAACAGAAC----- 1077  
 Qy 379 GlyAlaAspLysGlyGlyProThrLeuSerLysLeuGlySerAsnSerThrAlaGlnLeu 398  
 Db 1078 -----GAGAAGACCGGCGCGCTGTCCAAGCTGGGTCCAACTCGACGCGCGCAGCTC 1131  
 Qy 399 TyrProLysAspAspGlyGluGlyArgAlaAlaAlaValAlaMetProProAlaSerVal 418  
 Db 1132 CGGCCAAAGAGACGCGCGGAGGATGCGCGG-----GCATCGCGCGCGCGAGCTG 1185  
 Qy 419 MetThrArgLeuIleLeuMetValTrpArgLysLeuIleArgAsnProAsnThrTyr 438  
 Db 1186 ATGACGAGGCTCATCTGATCGTGTGGAGAGCTGATCAGGAACCCCAACACTTAC 1245  
 Qy 439 SerSerLeuGlyValValTrpSerLeuValSerTyrArgTyrGlyIleGluMetPro 458  
 Db 1246 TCCAGCGCTCTTGGTGTCTATCTGGTCCCTCGTCTCTCAAGTGGGGATTGAGATGCCA 1305

Qy 459 AlaIleIleAlaArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPhe 478  
 Db 1306 GCATCATCGCCCGGTGCAATTCGATCTGTGATGCAGGCTTGGAAATGGCCATGTTTC 1365  
 Qy 479 SerLeuGlyLeuPheMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAla 498  
 Db 1366 AGCTAGGATTTTCATGCAATTCGAGCCACGATCATTCGCTGGGAACCTCCCTTGTCT 1425  
 Qy 499 AlaIleAlaMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaAlaSerIle 518  
 Db 1426 TGTATGCCATGGCGTCAAGTTCCTCGTGGTCCGCGCTCATGGCTGCGCGCTCCATC 1485  
 Qy 519 AlaValGlyLeuArgGlyValLeuLeuHisIleAlaIleVal----- 532  
 Db 1486 GCGGTCCGATTCGCGGGTGTCTTGCACATTCGCAATTCGATTCAGTAGCTGCAAGTTC 1545  
 Qy 532 ----- 532  
 Db 1546 AGCTTCTGCACAAATGATCTTAACTGCAAAAGATCTATCTGTAACCTCAACAAATTG 1605  
 Qy 533 -----GlnAlaAlaLeuProGlnGlyIleValProPheValPheAlaLysGlu 548  
 Db 1606 TCATTGAATTCCTCAGGCGCTCTTCTCAAGGAATCGTGCCTTGTGTTCGCAAGGAG 1665  
 Qy 549 TyrGlyValHisProAspIleLeuSerThrAlaTyrGlyProIleThrSerHisGlyPhe 568  
 Db 1666 TACAATGTTTCATCTTAACATCTGAGCACAGCGTATGATCTCTTACCATCTTTCTTTT 1725  
 RESULT 4  
 AX652923 1686 bp DNA linear PAT 22-MAR-2003  
 LOCUS AX652923  
 DEFINITION Sequence 2793 from Patent WO03000898.  
 ACCESSION AX652923  
 VERSION AX652923.1 GI:29155737  
 KEYWORDS  
 SOURCE Oryza sativa  
 ORGANISM Oryza sativa  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.  
 REFERENCE 1  
 AUTHORS Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,  
 Katagiri, F., Quan, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.  
 TITLE Plant genes involved in defense against pathogens  
 JOURNAL Patent: WO 03000898-A 2793 03-JAN-2003;  
 SYNGENTA PARTICIPATIONS AG (CH)  
 FEATURES  
 source 1.1686  
 /organism="Oryza sativa"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:4530"  
 ORIGIN  
 Alignment Scores: 3.32e-105 Length: 1686  
 Pred. No.: 2020.50 Matches: 442  
 Score: 2020.50  
 Percent Similarity: 75.69% Conservativeness: 22  
 Best Local Similarity: 72.10% Mismatches: 39  
 Query Match: 69.43% Indels: 111  
 DB: Gaps: 16  
 US-10-030-884-14 (1-573) x AX652923 (1-1686)  
 Qy 1 MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValProLeuTyrValAla 20  
 Db 1 ATGATCAGCGTGGTGGACCTGTACCAGCTCTGACGCGGTGTGCGGTGTGACGTGGCG 60  
 Qy 21 MetThrLeuAlaTyrGlySerValArgTyrTrpArgIlePheThrProAspGlnCysSer 40  
 Db 61 ATCAGCGTGGGTACGCTCCGCGGTGGTGGCATCTTCTCCCGGACCATGTCCTCC 120  
 Qy 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60

Db 121 GGCATCAACCGCTTGGTCCGCTCTTCCGCTCCGCTCTCTCTCTCCATCTCATCTCC 180  
Qy 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80  
Db 181 ACCAACACCCCTTCCGCTCAACCTCCGCTTCTCTCCGCGCGCACACGCTCCAGAGCTC 240  
Qy 81 AlaValLeuAlaLeuLeuAla-----LeuAlaSerArgGlyLeuSerSerPro 96  
Db 241 ATCGTCTCCGCTCTCTCCGCTCTGGTCCGCTCTCTCCGCGCGGG----- 288  
Qy 97 ArgAlaLeuGlyLeuAspTrpSerIleThrLeuPheSerLeuSer-ThrLeuProAsnTh 116  
Db 289 -----TCCCTCGACTGGCTCATCACCTCTCTCTCCCTCTCCAACTTCCCAAMAC 339  
Qy 116 rleuValMetGlyIleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLe 136  
Db 340 CCTCGTCATGGCATCCCGCTCTCAAGGAT----- 372  
Qy 136 uMetValGlnValValLeuGlnCysIleIleTyrThrLeuMetLeuPheLeuPh 156  
Db 373 -----TGCATCATCTGGTACACGCTCATGCTGTTCTCTT 407  
Qy 156 eGluTyrArgAlaAlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSe 176  
Db 408 CGAGTACCGTCCGCGCGCTTCTTGTATGAGGAGGAGTCCCGGAC-----ACGCGCGCTC 464  
Qy 176 rIleValSerPheArgValAspSerValValSerLeuAlaArgGlyAsp----- 193  
Db 465 CATCGTCTCTCCGCTGGTATCCGAGCTGCTCGCTCGCGCGCGCTGGTGGAGC 524  
Qy 194 -ValGluLeuGluAlaGluProAspGlyValAlaGlyAlaValSerSerArgG 213  
Db 525 CCGGAGCTGCGAGCGGAG-----GCGAGGT 551  
Qy 213 yGlyAspAlaGlyArgValArgValThrValArgLysSerThrSerSerArgSerGluAl 233  
Db 552 CGGGGACGATGCACAGATCCGAGTCCCGTCCGCGAGTCCGAGCTCGCGCTCCGAGGC 611  
Qy 233 alaCysSerHis-----SerHisSerGlnThrMetGlnProArgValSerAsnLe 250  
Db 612 GCGGTGCTCGCACGCGAGCTCACTCGCAGTCCATGCGCGCGCGCTCCCAAGCT 671  
Qy 250 uSerGlyValGluIleTyrSerLeuGlnSerSerArgAsnProThrProArgGlySerSe 270  
Db 672 CTCGCGGCTGAGATTACTCTGCTGAGTCTGCTGCGGACCCGACGCGCGGCTCCAG 731  
Qy 270 rPheAsnHisAlaAspPheAsnIleValGlyAlaAlaAlaLysGlyGlyGlyAl 290  
Db 732 CTTCAACACCGCGAGTCTTCAACATCGTC-----GGCAACGCGCA 773  
Qy 290 aAlaGlyAspGluGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 310  
Db 774 GCACGCGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 827  
Qy 310 oGlnAlaValAlaValProAlaLysArgLysAspLeuHisMetLeuValTrpSerSerSe 330  
Db 828 G-----GTGTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 875  
Qy 330 rAlaSerProValSerGluArgAla-----AlaValHisValPh 343  
Db 876 CGGCTCGCGGCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 935  
Qy 343 eGlyAlaGlyGlyAlaAspHisAlaAspValLeuAlaLysGlyGlyGlyGlyGlyGlyGly 363  
Db 936 CGGCGGTGGTGGCGGCGGACGACGCGGAC-----GCCAAAGGTGCTCAGGCTTATGATGA 989  
Qy 363 uTyrGlyArgAspAspTyrSerSerArgThrLysAsnGlySerGlyGlyAlaAspLysG 383  
Db 990 G-----TACAGCTTCGGGAAACAGAAAT-----GAGAGGA 1019  
Qy 383 yGlyProThrLeuSerLysLeuGlySerAsnSerThrAlaGlnLeuTyrProLysAspAs 403

Db 1020 CGGCGCGCACACTGTCTCAAGCTCGGGTCCAACTCGACGCGCGAGCTCCGGCCAAAGACGA 1079  
Qy 403 pGlyGluGlyArgAlaAlaAlaValAlaMetProAlaSerValMetThrArgLeuI 423  
Db 1080 CGGCGAGGGGAGGGGGCA-----GCGATCGCGCGCGAGCGGTGATGACGAGGCTCAT 1133  
Qy 423 eLeuIleMetValTrpArgLysLeuIleArgAsnProAsnThrTyrSerSerLeuIleG 443  
Db 1134 ACTGATCATGGTCTGGAGGAGCTGATCAGGAACCCAAACACTTACTCCAGGCTCTCTCG 1193  
Qy 443 yValValTrpSerLeuValSerTyrArgTrpGlyIleGluMetProAlaIleAlaAla 463  
Db 1194 TGTCTATCTGCT 1253  
Qy 463 gSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeuGlyLeuPh 483  
Db 1254 GTCGATTTCTGATCTCTTCTGATGTCAGGGCTCGGAAATGGCCATGTTTCAGCTAGGATGTT 1313  
Qy 483 eMetAlaLeuGlnProArgIleAlaCysGlyAsnLysLeuAlaAlaIleAlaMetG 503  
Db 1314 CATGGATTTGACCCACGAGTCTATGCTTGGGAATCCCTTGTCTGATGATGATGATGATG 1373  
Qy 503 yValArgPheValAlaGlyProAlaValMetAlaAlaAlaSerIleAlaValGlyLeuAr 523  
Db 1374 CGTCAGGTTCTCTGTTGGTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 1433  
Qy 523 gGlyValLeuLeuHisIleAlaIleVal----- 532  
Db 1434 CGGGGTCTCTCTGACATTCGCAATTCGCAATTCGCAATTCGCAATTCGCAATTCGCAATTC 1493  
Qy 533 -----G 533  
Db 1494 AAATGTCATCTTAATGTCAAAAGATCTATCTGTAATCTCAACATTCGATTCATTCATTC 1553  
Qy 533 nAlaAlaLeuProGlnGlyIleValProPheValPheAlaLysGlyTyrGlyValHisPr 553  
Db 1554 GCGCGCTTCTTCTTCAAGGAATCGTCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1613  
Qy 553 oAspIleLeuSerThrAlaTyrGlyProIleThrSer 565  
Db 1614 TAACATTTCTGACGACAGCGTATGATCTCTTTACCACG 1650  
RESULT 5  
CNS08CDY 91883 bp DNA linear PLN 21-NOV-2003  
LOCUS Oryza sativa chromosome 12, . BAC OJ1575.G05 of library Monsanto  
DEFINITION from chromosome 12 of cultivar Nipponbare of ssp. japonica of Oryza  
sativa (rice), complete sequence.  
ACCESSION BX000508.1 GI:25900581  
VERSION  
KEYWORDS HTG.  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Eriactoidae; Oryzaceae; Oryza.  
REFERENCE 1 (bases 1 to 91883)  
AUTHORS Choinshe, N., Orjeda, G., Cattolico, L., Demange, N., Wincker, P.,  
Segurens, B., Pelletier, E., Scarpelli, C., Salanoubat, M.,  
Weissenbach, J. and Quetier, F.  
Oryza sativa chromosome 12 sequencing  
2 (bases 1 to 91883)  
Genoscope.  
Direct Submission  
Submitted (20-NOV-2003) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
Center code: GS  
Center code: GS  
Web site: http://www.genoscope.cns.fr/  
Contact: Seqref@genoscope.cns.fr  
-----



The following sequence is oriented from the T7 to the SP6 end. The nucleotide sequence of this BAC clone was generated by combining Monsanto and Genoscope sequencing data.

Upstream BAC (overlapping the T7 end) : OSUNBa0009F13 (AC-BX000498)

Downstream BAC (overlapping the SP6 end) : OJ1239\_F07 (AC-BX000504)

----- Finishing boundaries -----

FINISHED SEGMENT STARTS AT BASE 28300

FINISHED SEGMENT ENDS AT BASE 69200

## FEATURES

source Location/Qualifiers

1. 91883

/organism="Oryza sativa (japonica cultivar-group)"

/mol\_type="genomic DNA"

/cultivar="Nipponbare"

/sub\_species="japonica"

/db\_xref="taxon:39947"

/chromosome="12"

/clone="OJ1575\_G05"

/clone\_lib="Monsanto"

## ORIGIN

## Alignment Scores:

Pred. No.: 1.23e-100 Length: 91883

Score: 1977.00 Matches: 458

Percent Similarity: 59.34% Conservative: 25

Best Local Similarity: 56.27% Mismatches: 39

Query Match: 67.94% Indels: 292

DB: 8 Gaps: 19

US-10-030-884-14 (1-573) x CNS08CDY (1-91883)

Qy 1 MetileThrAlaLeuAspLeuTyHisValLeuThrAlaValProLeuTyValAla 20

Db 32692 ATGATCACGGTGGTACCGCTCTGACCGCGGTGGTGGCGTGTGACGTGGCG 32751

Qy 21 MetThrLeuAlaTyGlySerValArgTTPArgilePheThrProAspGlnCysSer 40

Db 32752 ATGACGTGGGTACCGCTCGTCCGGTGGGCGCATCTTCTCCCGACCAAGTGTCC 32811

Qy 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60

Db 32812 GGCATCAACCGCTTCGTGGCGCTCTTCGCGTCCCGTCTCTCTCTCCACTTCATCTCC 32871

Qy 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80

Db 32872 ACCAACACCCCTTCGCGCATGAACCTCCGCTTCCTCCGCGCGACACGCTCCAGAGCTC 32931

Qy 81 AlaValLeuAlaLeuLeuAla-----LeuAlaSerArgGlyLeuSerSerPro 96

Db 32932 ATCGTCTCGCCCTCTCTCGCGCTCTGGTGGCGCTCTCCGCGCGCGG----- 32979

Qy 97 ArgAlaLeuGlyLeuAspTTPSerIleThrLeuPheSerLeuSerThrLeuProAsnThr 116

Db 32980 -----TCCCTCGACTGGCTCATACCCCTCTCTCTCCCTCTCCACCCCTCCCAACACC 33030

Qy 117 LeuValMetGlyIleProLeuArgGlyMetTyrglyAlaSerAla----- 133

Db 33031 CTCGTATGGGCATCCCGCTCTCAAGGGATGACCGCGCGCGCGCGCGCGCGCGCG 33090

Qy 134 -----GlyThrLeuMetValGlnValValValLeuGlnCysIleIleTTPTyThr 150

Db 33091 GCCGACTCCGCGAGCTCATGTGCAGATCGTCTGTCTCCAGTGCATCATCTGTGTACAG 33150

Qy 151 LeuMetLeuPheLeuPheGlnTyArgAlaAlaArgAlaLeuValLeuAspGlnPhePro 170

Db 33151 CTCATGCTGTTCCTCTTCGATGACCGTGGCGCGCGCTCTTGTTCATGAGCAGTTCCTCG 33210

Qy 171 AspGlyAlaAlaAlaSerIleValSerPheArgValAspSerAspValSerLeuAla 190

Db 33211 GAC---ACCGCGCTCTCATGCTCTCTCTCCGGTGGATTCGAGCTGCTCTGCTCGCC 33267

Qy 191 ArgGlyAsp-----ValGlnLeuGlnAlaLupProAspGlyValAlaGlyAlaGly 207

Db 33268 GCGCGCGCTGGTGGAGCCGCGGAGCTGCAGCGGAG----- 33303

Qy 208 AlaValSerSerArgGlyGlyAspAlaGlyArgValArgValThrValArgLysSerThr 227

Db 33304 -----GCGAGCGTGGGAGCATGCGAAGTTCGAGTTCGCGTGGCAAGTCGACG 33354

Qy 228 SerSerArgSerGluAlaAlaCysSerHis-----SerHisSerGlnThrMetGln 244

Db 33355 AGCTCGCGCTCGAGCGCGCGTCTCGACGGGAGCGAGTCACTCGCAGTCCATGCGAG 33414

Qy 245 ProArgValSerAsnLeuSerGlyValGlnLeuTySerLeuGlnSerSerArgAsnPro 264

Db 33415 CGCGCGCTCTCAACCTCTCGCGCTCGAGATTTACTCGCTGCAGTCTCGCGAACCAG 33474

Qy 265 ThrProArgGlySerSerPheAsnHisAlaAspPhePheAsnIleValGlyAlaAla 284

Db 33475 AGCGCGCGGCTCCAGCTTCAACCCCGCGAGTCTTCAACATCGTC----- 33522

Qy 285 LysGlyGlyGlyAlaAlaGlyAspGlnLysGlyAlaCysGlyGlyGlyGly 304

Db 33523 -----GCCACCGCAGCACCGCGCAGGAGAGGGCGCC-----GCCGCTGGCGGG 33570

Qy 305 GlyHisSerProGlnProGlnAlaValProAlaLysArgLysAspLeuHisMet 324

Db 33571 GGCCACTCGCGCAGCG-----GTGGTGGGAGAGAGAGAGACCTGCACATG 33618

Qy 325 LeuValTTPSerSerSerAlaSerProValSerGluArgAla----- 338

Db 33619 TTCGTGTGGAGCTCAAGCGCTCGCGGTGTGGAGCGCGCGCGCGCGCGCTGCC 33678

Qy 339 ---AlaValHisValPheGlyAlaGlyAlaAspHisAlaAsp----- 352

Db 33679 GCGCGGTGTCAGCTTTCGCGCGTGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 33738

Qy 352 -----ValLeu-Ala----- 355

Db 33739 ACCTTTTATCATCTTTTGGTGCATTTGACTTCCCTTTTGTGGCGCGGTATT 33798

Qy 353 -----ValLeu-Ala----- 355

Db 33799 ACATGTTGTTTCTTCTGCACATTAATCACTCTGTCTGTCTGTGCTCTTGACCATCCGAT 33858

Qy 355 ----- 355

Db 33859 CAGTATATCAACGGTCCAGATCACACTGGTCTGTGATTGTTGGAGAACTCAACTCG 33918

Qy 355 ----- 355

Db 33919 TATACTGATTAAATACCACTTCTTGAAGTGAATCTTTTTCACAAACTTACCTTCTTGA 33978

Qy 355 ----- 355

Db 33979 TTGTTAAATTCATGTAGCTTTTCTGATCTTAAACAGTGACATGACTGTTAAGTAGATTG 34038

Qy 355 ----- 355

Db 34039 AATGTGACAGTTAAGACAAATCTTAAACAGTAAACTACTGATTTCACAAAGAAAATGT 34098

Qy 356 -----LysGlyAlaGlnAlaTyArgGlnTyrglyArgAspAspTySerSe 371

Db 34099 CTTTGATGATGACACAGGTGCTCAGCGCTTATGATGAG-----TACAGCTT 34143

Qy 371 rArgThrLysAsnGlySerGlyGlyAlaAspLysGlyGlyProThrLeuSerLysLeuG 391

Db 34144 CGGGAACAAGAT-----GAGAAGCAGCGCGCGCGACACTGTCTGAAGCTGGG 34188

Qy 391 ySerAsnSerThrAlaGlnLeuTyProLysAspAspGlyGlyGlyArgAlaAlaAlaVa 411

Db 34189 GTCCAACTCGAGCGCGCAGCTCCGCCAAAGACCGCGCGAGGGAGGGCGCGCA----- 34243

Qy 411 lAlaMetProAlaSerValMetThrArgLeuIleLeuIleMetValTTPArgLysLe 431

Db 34244 -GCGATGCCCGCGCGAGCGCTGATGACGAGGCTCATATCATGATCATGCTGTGAGGAAGCT 34302



Qy 431 uileArganProAntThrTySerSerLeuIleGlyValValTrpSerLeuValSerTy 451  
 Db 34303 GATCAGGAACCAAACTTACTCCAGCCTCTCGGTGTCATCTGGTCCCTCGTCTCCTA 34362  
 Qy 451 r----- 451  
 Db 34363 CAGGTCAATTAAGCAACATTTTACTCACACTTTTTCACATGGCCATGTTTCATCATGCTC 34422  
 Qy 452 -----ArgTrpG1 454  
 Db 34423 ATCTCTGTTGTGCATGTGTTTAACTCTGTGTAACAATGTGATGTAATTCAGGTGGG 34482  
 Qy 454 VileGluMetProAlaIleAlaAArgSerIleSerIleLeuSerAaspAlaGlyLeuG1 474  
 Db 34483 AATTGAGATGCGGTATCATCGCCCGGTGATTCGATTCGATTCCTTCAGATGACAGGCTCG 34542  
 Qy 474 yMetAlaMetPheSerLeu----- 480  
 Db 34543 AATGGCCATGTTTCAGCCTAGGTTTGTGTGCATCTCTGATACATGCTGTCATTTGCTGTA 34602  
 Qy 481 -----GlyLeuP 483  
 Db 34603 TCTTACATGGTTCATGCTGATGCAATGACTGACTGATGTTCTGTTTCAGAGATGT 34662  
 Qy 483 heMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAlaIleAlaMetG 503  
 Db 34663 TCATGGCATTTGAGCCAGGATCATGCTGTGGAACTCCCTTCTGTTATGATGCTATGG 34722  
 Qy 503 yValArgPheValAlaGlyProAlaValMetAlaAlaSerIleAlaValGlyLeuA 523  
 Db 34723 CGGTCAAGTTCCTGTTGCTGCGGTCATGGCTGCGGCTTCATTCGCGCTCGACTTC 34782  
 Qy 523 rGlyValLeuLeuHisIleAlaIleVal----- 532  
 Db 34783 GGGGGTGCTTCTGCACATTTGCTGTTTTCAGGTAAGCTGCAAGTTCAGGCTTCTGCACC 34842  
 Qy 533 -----G 533  
 Db 34843 AAAATGCATCTTAATCTGCAAAAGAAATCTATCGTAACCTCAACAATGTTCATGACTCTC 34902  
 Qy 533 InAlaAlaLeuProGlnGlyIleValProPheValPheAlaLysGlyTrpGlyValHisP 553  
 Db 34903 AGCGCGCTTCTCTCAAGGAATCGTGGCTTCTGTTTGGCAAGAGTACAATGTTTCATC 34962  
 Qy 553 roAspIleLeuSerThrAlaTrpGlyProIleThrSer 565  
 Db 34963 CTAACATTTCTGACGACAGCGTATGATCTCTTTTACCACC 35000

RESULT 6  
 AC123528/c 141025 bp DNA linear HTG 29-MAY-2002  
 LOCUS Oryza sativa chromosome 11 clone OSUNBa0078E03, \*\*\* SEQUENCING IN  
 DEFINITION PROGRESS \*\*\*, 3 ordered pieces.  
 ACCESSION AC123528  
 VERSION AC123528.1 GI:21240714  
 KEYWORDS HTG; HTGS; PHASE2.  
 SOURCE Oryza sativa  
 ORGANISM Oryza sativa  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.  
 REFERENCE 1 (bases 1 to 141025)  
 AUTHORS Linton, E.W., Tanyolac, B., Young, S., Kavchok, S., Keizer, G.,  
 Bronzino, A., and Messing, J.  
 TITLE Unpublished  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 141025)  
 AUTHORS Linton, E.W., Tanyolac, B., Young, S., Kavchok, S., Keizer, G.,  
 Bronzino, A., and Messing, J.  
 TITLE Direct Submission  
 JOURNAL Submitted (29-MAY-2002) The Plant Genome Initiative at Rutgers -  
 Waksman Institute, Rutgers, The State University of New Jersey, 190

Prelinghuysen Road, Piscataway, New Jersey 08854, USA  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 3 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* provided by the submitter.

\* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.  
 \* 1 12797 12896: contig of 12796 bp in length  
 \* 12797 12896: gap of unknown length  
 \* 109309 109408: contig of 96412 bp in length  
 \* 109309 109408: gap of unknown length  
 \* 109409 141025: contig of 31617 bp in length.

FEATURES  
 source  
 1..141025  
 /organism="Oryza sativa"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:4530"  
 /chromosome="11"  
 /clone="OSUNBa0078E03"

## ORIGIN

## Alignment Scores:

Pred. No.: 2,01e-99 Length: 141025  
 Score: 1959.50 Matches: 456  
 Percent Similarity: 58.92% Conservative: 26  
 Best Local Similarity: 55.75% Mismatches: 43  
 Query Match: 67.34% Indels: 294  
 DB: 2 Gaps: 19

US-10-030-884-14 (1-573) x AC123528 (1-141025)

Qy 1 MetIleThrAlaLeuAspLeuTyHisValLeuThrAlaValProLeuTyValAla 20  
 Db 62955 ATGATCAGGTGTGTGACCTGTACCACTGCTGACGCGGTGTGCGTGTACGTGGC 62896  
 Qy 21 MetThrLeuAlaTrpGlySerValArgTrpTrpArgIlePheThrProAspGlnCysSer 40  
 Db 62895 ATGACGTGGCGTACCCCTCCGTCGGTGTGGCGCATTTCTCCCGACCACTGCTCC 62836  
 Qy 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
 Db 62835 GGCATCAACCGCTTCGTGCGCTCTTCGCGCTCTCTCTCTCTCTCTCTCTCTCTCC 62776  
 Qy 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80  
 Db 62775 ACCAACAACCCCTTCGCCATGAACCTCGCTTCTCGCGCGCACACGCTCCAGAACTC 62716  
 Qy 81 AlaValLeuAlaLeuLeuAla-----LeuAlaSerArgGlyLeuSerSerPro 96  
 Db 62715 ATGCTCTCGCCCTCTCTCGCCCTCTGTCGCGCTCTCTCGCGCGCGG----- 62668  
 Qy 97 ArgAlaLeuGlyLeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThr 116  
 Db 62667 -----TCACTCGACTGGCTCATCACCTCTCTCTCTCTCTCTCTCTCTCTCTCC 62617  
 Qy 117 LeuValMetGlyIleProLeuLeuArgGlyMetTyTrpGlyAlaSerSer-----Ala 133  
 Db 62616 CTGTCAATGGGATCCCGTGTCTCAAGGGAGTGTACGCCCGCGCGCGCGCGCGCGCTCC 62557  
 Qy 134 GlyThrLeuMetValGlnValValValValLeuGlnCysIleIleTrpTrpThrLeuMetLeu 153  
 Db 62556 GGCAGCCTCATGTTGTCAGATCGTCTGCTCCAGTGTATCATCTGTGTACACGCTCATGCTG 62497  
 Qy 154 PheLeuPheGluTyTrpArgAlaAlaArgAlaLeuValLeuAspGlnPheProAspGlyAla 173  
 Db 62496 TTCTCTTTCAGTACCGCGCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 62440  
 Qy 174 AlaAlaSerIleValSerPheArgValAspSerAspValValSerLeuAlaArgGlyAsp 193  
 Db 62439 GCGGCTCCATCT 62380

QY 194 -----ValGluLeuGluAlaGluProAspGlyValAlaGlyAlaValSer 210  
 DB 62379 GGTGGAGCGCGGAGCTGACGGCGAG----- 62353  
 QY 211 SerArgGlyGlyAspAlaGlyArgValArgValThrValArgLysSerThrSerSerArg 230  
 DB 62352 CGGGAGGTTCGGGGACGATGGCAGGATGCGGGTCACCGTGGCCAAAGTCGACAGGTCGCGC 62293  
 QY 231 SerGluAlaAlaCysSerHis-----SerHisSerGlnThrMetGlnProArgVal 247  
 DB 62292 TCCGAGGGGGGTGCTCCGACGGGACGAGTCACACTCGCAGTCGACGCGCGGGTC 62233  
 QY 248 SerAsnLeuSerGlyValGluLeuThrSerLeuGlnSerSerArgAsnProThrProArg 267  
 DB 62232 TCCAACTCTCGGGTGCAGATTTACTCGTGCAGTCGTGCGGAAACCCGACGCGCGT 62173  
 QY 268 GlySerSerPheAsnHisAlaAspPhePheAsnLeuValGlyAlaAlaAlaLysGlyGly 287  
 DB 62172 GGCTCAGCTTCACACGCGCGAGTCTTCAACATCGTC-----GGC 62131  
 QY 288 GlyGlyAlaAlaGlyAspGluGluLysGlyAlaCysGlyGlyGlyGlyHisSer 307  
 DB 62130 AACGGCAAGCGGCGACGAGGAGAGGGCGC-----GCCGTGGCGGGCGCACTCG 62077  
 QY 308 ProGlnProGlnAlaValAlaValProAlaLysArgLysAspLeuHisMetLeuValTrp 327  
 DB 62076 CGCGAGCGG-----GTGGTGGGAGAGAGGACCTGACATGTTCTGTGTGG 62029  
 QY 328 SerSerSerAlaSerProValSerGluArgAla-----AlaValHis 341  
 DB 62028 AGCTCAAGCGCTCCCGGTGTCGAGCGCGCGCGCGAGCGAGCTGGCGGCGGCAC 61969  
 QY 342 ValPheGlyAlaGlyGlyAlaAspHisAlaAsp----- 352  
 DB 61968 GTCTTCGGCGGTGTGGCGCGCACCGACCGCGCGCAAGGTGACGACACCTTTTGGAT 61909  
 QY 352 ----- 352  
 DB 61908 CATCTTTTCGTGCATTTGACTTCCCTTTTGTGTCGGCGGTGTAATACACTGTTGTT 61849  
 QY 352 ----- 352  
 DB 61848 GCTTCTGGCAATTAATCACTCTTGTCTGTCTTGCATCCGATCCGATAGAGATCAA 61789  
 QY 352 ----- 352  
 DB 61788 CGGTCCAGATCACACTAGTCTGCTCGTTGTTGGAGAACTCGACAGCCCATACATATTA 61729  
 QY 352 ----- 352  
 DB 61728 ATACTACATCTTGAAGTGATTTTTTTTTCACAACTTACTTTCTTGAATTTGTTACTTTT 61669  
 QY 352 ----- 352  
 DB 61668 CTGATCTTAACAGTGACATGACTGTTAACTGTAAGTATAGATGATGATGATGAGTTGA 61609  
 QY 353 -----ValLeuAlaLys----- 356  
 DB 61608 ATTGTACTTTAAGACAAAATCTTAACAGTAACTACTGATTTGACAAAGAGATGTC 61549  
 QY 357 -----GlyAlaGlnAlaLysAspGluTyrGlyArgAspAspTyrSerSer 371  
 DB 61548 TTTGATGGATGACAGGAGCTCAGGCTATGATGAG-----TACAGTTTC 61504  
 QY 372 ArgThrLysAsnGlySerGlyGlyAlaAspLysGlyGlyProThrLeuSerLysLeuGly 391  
 DB 61503 GGAACACAGAAC-----GAGAAGGAGCGGCGGACGCTCTCCAGCTGGG 61459  
 QY 392 SerAsnSerThrAlaGlnLeuTyrProLysAspAspGlyGlyGlyArgAlaAlaAlaVal 411  
 DB 61458 TCCAACTCGACGCGCAGCTCCGGGCAAGAGGACGAGCGGCGGATGCGCGC----- 61405

QY 412 AlaMetProProAlaSerValMetThrArgLeuLeuLeuMetValTrpArgLysLeu 431  
 DB 61404 GCGATCCGCGCGGAGCGTGTGATGACAGGCTCATCTGATCATGTTTGGAGGAAGCTG 61345  
 QY 432 IleArgAsnProAsnThrTyrSerSerLeuIleGlyValValTrpSerLeuValSerTyr 451  
 DB 61344 ATCAGGAACCCCAACACTTACTCCAGCCCTCTTGGTGTGATCTGGTCCCTCGCTCTTAC 61285  
 QY 452 Arg----- 452  
 DB 61284 AG-GTTAATTAAGCAAAACATTTTACTCACAATTTTTCACATTGACATGTTTCATTAGCC 61226  
 QY 453 -----TrpGly 454  
 DB 61225 ATCTCTGTTGTCGCAATTTGTTGATCTGTGAAACAATGTGATGTAATTCAGTGGGG 61166  
 QY 455 IleGluMetProAlaIleIleAlaArgSerIleSerIleLeuSerAspAlaGlyLeuGly 474  
 DB 61165 ATTGAGATGCCAGCATCATCGCCCGTCTGATTCGATCTCTGCAGATCGAGGCTTGA 61106  
 QY 475 MetAlaMetPheSerLeu----- 480  
 DB 61105 ATGGCCATGTTTCAGCCCTAGGTTTGTGTTTCATCTCTGTGATCATATGATAAATTTGCTGA 61046  
 QY 481 -----G 481  
 DB 61045 ATTTTACATGCGTTCATGATGAGACTGCAATAACTGAATTTGTTTTCGAGG 60986  
 QY 481 YLeuPheMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAlaIleAl 501  
 DB 60985 ATTGTTTCATGGCATTCGACCCAGGATCATTCGCTGGGAACTCCCTTGTGTTATGC 60926  
 QY 501 aMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaAspIleAlaValG 521  
 DB 60925 CATGGCGTCAGTCTCTCGTGGTCCGCGCTGATGCTGCGCTCCATCGCCCTCGG 60866  
 QY 521 YLeuArgGlyValLeuLeuHisIleAlaIleVal----- 532  
 DB 60865 ACTTCGGGGTCTTCTGCACATTCGCAATGTTTTCAGTAAAGTCAAGTTCAGGCTTCT 60806  
 QY 532 ----- 532  
 DB 60805 GCACCAAAATGCATCTTAACATGCAAAAGAACTATCTGTAACCAACAATTTGATGAA 60746  
 QY 533 -----GlnAlaAlaLeuProGlnGlyIleValProPheValPheAlaLysGlyTyrGlyVa 551  
 DB 60745 TTCTCAGCGCGCTCTTCTCCTCAGGAATCGTCCCTTGTGTTGTTGCCAAGAGTACAATGT 60686  
 QY 551 LHsProAspIleLeuSerThrAlaTyrGlyProIleThrSerHisGlyPhe 568  
 DB 60685 TCATCTCAATCTTGACGACGCGTATGATCTCTTTACCATCTTTTCTTTT 60634  
 RESULT 7  
 AF515435 2297 bp mRNA linear PLN 20-AUG-2003  
 LOCUS Populus tremula x Populus tremuloides PIN1-like auxin transport  
 DEFINITION protein (pin2) mRNA, complete cds.  
 ACCESSION AF515435  
 VERSION AF515435.1 GI:21435939  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Populus tremula x Populus tremuloides  
 Populus tremula x Populus tremuloides  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids 1; Malpighiales; Salicaceae; Salicaceae; Populus.  
 1 (bases 1 to 2297)  
 Schradler, J., Baba, K., May, S.T., Palme, K., Bennett, M., Bhalerao, R.P.  
 and Sandberg, G.  
 Polar auxin transport in the wood-forming tissues of hybrid aspen  
 is under simultaneous control of developmental and environmental  
 signals  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (17), 10096-10101 (2003)  
 PUBMED 12909722

REFERENCE 2 (bases 1 to 2297)  
 AUTHORS Baba, K., Schrader, J., Palme, K., Bhalerao, R. P. and Sandberg, G.  
 TITLE A family of PIN1 like auxin transporters in hybrid aspen  
 JOURNAL Unpublished  
 REFERENCE 3 (bases 1 to 2297)  
 AUTHORS Baba, K., Schrader, J., Palme, K., Bhalerao, R. P. and Sandberg, G.  
 TITLE Direct Submission  
 JOURNAL Submitted (24-MAY-2002) Forest Genetics and Plant Physiology,  
 Swedish University of Agricultural Sciences, Umea 90183, Sweden  
 FEATURES  
 source Location/Qualifiers  
 1..2297 /organism="Populus tremula x Populus tremuloides"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:47664"  
 1..2297  
 /gene="pin2"  
 339..2105  
 /gene="pin2"  
 /codon\_start=1  
 /product="PIN1-like auxin transport protein"  
 /protein\_id="AA54034.1"  
 /db\_xref="GI:21435940"  
 /translation="MISIVDLVHLTVAVPLVAVMLAVGSKVKKWKFSPDQCSGINSR  
 FVALFAPLPSLHFRPISNNPYAMNGLGFIADTLQKILVILVLAWSRASSRSGLEMSI  
 TLFSSSLPTLVMLGKMGNSGSLMVOIVLQCIITWITLMLFLFETRGARIL  
 IGEQPDITAGSIISPRVDIILSDGREPLQTEAEVDEGKLHVTVKSTSRSRVFS  
 HMSGNLGSLTPRPSNLTAELIYSLQSRNPTFRASFNHTDYFVSKNGKASPRH  
 SNFTLPQBSVSGGLVFGNVRANGSAVPAFPNAGIFSPGGKKKANGKNGKIDHMF  
 VNSSASPVSEGLHVFEGDYGNDLGGVAHKHVDYDFGRDSEFSGNRPNGNVDRDG  
 PVLKLGSSSTAEHLKPSKANGKPTWMPPTVMTKLILINVMKLLRNPNYSSLI  
 GLTWLSVFPKWNIPATIANISAILSNAGLGNMPSLGLFWALPRIACGNSIASF  
 AMGVRLETPGVNMAASFAVGURGLDRLRIAVQAAIPQGLVFPVFAKGINVHPDILST  
 GVIFGMLIALPITLVYVILLGL"

## ORIGIN

Alignment Scores:  
 Pred. No.: 8 51e-92 Length: 2297  
 Score: 1785.50 Matches: 384  
 Percent Similarity: 71.94% Conservative: 57  
 Best Local Similarity: 62.64% Mismatches: 89  
 Query Match: 61.36% Indels: 83  
 DB: 8 Gaps: 16

US-10-030-884-14 (1-573) x AF515435 (1-2297)

Qy 1 MetIleThraLeuAspLeuTyrHisValLeuThraValAlaValProLeuTyrValAla 20  
 Db 339 ATGATCATGATCGTAGACCTCTACCATGCTTACAGCTGTGGTACCCCTCTATGTGGCC 398  
 Qy 21 MetThraLeuAlaTyrGlySerValArgTyrTrpArgIlePheThrProAspGlnCysSer 40  
 Db 399 ATGATTTGGCTTATGTTCTGTGCAATGTTGGAAATCTTTAGCCCTGATCAATGTCA 458  
 Qy 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
 Db 459 GGGATCAACAGATTGTGTCTGTGTGCAGTACCTTGTCTTTCTTCTTCACTTCATTCC 518  
 Qy 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80  
 Db 519 TCCAAACACCCCTAGCTAAGAACTAGGGTTCACTGCTGCAGATACTCTTCAAAAAT 578  
 Qy 81 AlaValLeuAlaLeuLeuAlaLeuAlaSerArgGlyLeuSerSerProArgAlaLeuGly 100  
 Db 579 ATTGTCTTGGTGTATTAGTAATTGTGCCAGA-----GCTAGCTCCAGAGCC---TCC 629  
 Qy 101 LeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThrIleValMetGly 120  
 Db 630 CTTGAATGGTCCATTACTTGTGTTTCACTCTCTAGTCTCCCAACACTCTTGTATGGGC 689  
 Qy 121 IleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeuMetValGlnVal 140  
 Db 690 ATCCCTTGTGTAAGGGAATGATGGA---AATCTCTCAGGGAGTCTCATGTGCCAATA 746

Qy 141 ValValLeuGlnCysIleIleTyrThrLeuMetLeuPheLeuPheGluTyrArgAla 160  
 Db 747 GTTGTCTTCAATGCATATATATGTCACATTCATGCTGCTTCTTGTGATATAGAGGA 806  
 Qy 161 AlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSerIleValSerPhe 180  
 Db 807 GCTAGAAATCTGATTGGTGAACAGATTCTCTGAC---ACAGCTGGCTCTATAATATCTTTC 863  
 Qy 181 ArgValAspSerAspValValSerLeu---AlaArgGlyAspValGluLeuGluAlaGlu 199  
 Db 864 AGGGTGTGATCTCATATCTTCTTTAGATGGTAGAGGCCATTCGAAACCGAGGCTGAA 923  
 Qy 200 ProAspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyClyAspAlaGlyArgVal 219  
 Db 924 GTC-----GGTGAAGATGGGAAGTTA 944  
 Qy 220 ArgValThrValArgLysSerThrSerSerArgSerGluAlaAlaCysSerHisSerHis 239  
 Db 945 CATGTTACTGTTCCGAATCACTAGTTCAGATCAGAGCTTTTCTCATATGTCCTCAT 1004  
 Qy 240 -----SerGlnThrMetGlnProArgValSerAsnLeuSerGlyValGluIle 255  
 Db 1005 GGTCTAAATCTGTGCTCTTCACTGACTCTCTAGACCATCTAATTTGACCAATGCAGAGATA 1064  
 Qy 256 TyrSerLeuGlnSerSerArgAsnProThrProArgGlySerSerPheAsnHisAlaAsp 275  
 Db 1065 TACTCTCTACATCTCCTCAAGAAATCCTCCCAAGAGCTTCGAGTTTAAACCATACCGAT 1124  
 Qy 276 PhePheAsnIleValGlyAlaAlaLysGlyGlyGlyAlaAlaGly----- 292  
 Db 1125 TTTTACTCCATGGTT-----AATGGCAAGAATCGAGTCCAGACACTCT 1169  
 Qy 293 -----AspGluGluLysGlyAlaCysGlyGlyGlyGlyGlyGly 305  
 Db 1170 AACTTCACTAATTGCAATTCGACGAAGAAGTGA-----GCACTTGGGTATTGGT 1223  
 Qy 306 HisSerProGlnProGlnAlaValAlaValProAla----- 317  
 Db 1224 AATGTTCCAGAGCAATGGAAGTCTTATCTCTCCACCAATGCTGGAATCTTTCT 1283  
 Qy 318 -----LysArgLysAspLeuHisMetLeuVal 326  
 Db 1284 CTGGGGGTAAAAAGAGCGCAATGGGGTGGAGAAATGGCAAGGATTTGCATATGTTGTT 1343  
 Qy 327 TrpSerSerAlaSerProValSerGluArgAlaAlaValHisValPheGlyAlaGly 346  
 Db 1344 TGGAGTTCAAGTGTTCACCAAGTATCTGAA---GGTGGCTACATGCTTTTAAGGGAGG 1400  
 Qy 347 Gly-----AlaAspHisAlaAspValLeuAlaLysGlyAla 358  
 Db 1401 GATTATGGCAATGACCTCGGTGGGTAGCTCACCATAAAGAT----- 1442  
 Qy 359 GlnAlaTyrAspGluTyrGlyValArgAspAspTyrSerSerArgThrLysAsnGlySerGly 378  
 Db 1443 -----TATGATGATGTTGGTTCGAGATGAGTTCAGCTTGGAAATAGACCTGGACCTAAT 1496  
 Qy 379 GlyAlaAspLysGlyGlyProThrLeuSerLysLeuGlySerAsnSerThrAlaGlnLeu 398  
 Db 1497 GGAGTGGATCGGGATGCTCCAGTCTTTTAAAGCTGGTTCAAGCTCTACAGCGCGCTA 1556  
 Qy 399 TyrProLysAspAspGlyGluGlyArgAlaAlaValAlaMetProAlaSerVal 418  
 Db 1557 CACCCAAAGATGCTGCTAATGTTGAAGCCCAAGCCAACTGTCATGCTCTCTACAGCGCTT 1616  
 Qy 419 MetThrArgLeuIleLeuMetValTrpArgLysLeuIleArgAsnProAsnThrTyr 438  
 Db 1617 ATGACAAGACTCACTCTCATATGTTGGGAGAAAATTTATAGAAAATCCAAATACCTAT 1676  
 Qy 439 SerSerLeuIleGlyValValTrpSerLeuValSerTyrArgTrpGlyIleGluMetPro 458  
 Db 1677 TCCAGCTGATGGCCCTCCTGCTGCTCTTCTTTCATTCAGTGGAACTAATGATGCC 1736  
 Qy 459 AlaIleIleAlaArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPhe 478

```

Db 1737 GCGATCATTGCTAAATCAATAGCTATTCTAATCTGCTGCTGGATGGCCATGTTT 1796
Qy 479 SerLeuGlyLeuPheMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAla 498
Db 1797 AGTCTTGCTGTTGTCATGGCATTACAGCCTAGGATTATCGCTGTGGAACTCAATCGCT 1856
Qy 499 AlaIleAlaMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaSerIle 518
Db 1857 TCCCTTGCCATGGGTGTAGATTCTCAGAGTCTCTCAGTCATGCTGCTGCTGTTT 1916
Qy 519 AlaValGlyLeuArgGlyValLeuLeuHisIleIleAlaIleValGlnAlaLeuProGln 538
Db 1917 GCTCTTGCTCTACAGAGAGATCTCTCGCATTCCTCCATAGTCAGGAGCTCTTCCACAA 1976
Qy 539 GlyIleValProPheValPheAlaLysGlyIleValHisProAspIleLeuSerThr 558
Db 1977 GGGATTGTGCTTGTGCTTGTGCAAGATATATGTTTCATCTGACACTACTGAGCACT 2036
Qy 559 Ala-----TyrGly-----ProIleThr 564
Db 2037 GGAGTTATATTGGGATGCTAATGCTCTACCCATCACA 2075

RESULT 8
AF056027 2358 bp mRNA linear PLN 03-AUG-1998
LOCUS Oryza sativa auxin transport protein REH1 (REH1) mRNA, complete cds.
DEFINITION
ACCESSION AF056027
VERSION AF056027.1 GI:3377508
KEYWORDS
SOURCE Oryza sativa
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 2358)
Luschnig, C., Gaxiola, R.A., Grisafi, P. and Pink, G.R.
EIR1, a root-specific protein involved in auxin transport, is
required for gravitropism in Arabidopsis thaliana
Genes Dev. 12 (14), 2175-2187 (1998)
MEDLINE 98344010
PUBMED 9679062
REFERENCE
1 (bases 1 to 2358)
Luschnig, C., Grisafi, P. and Pink, G.R.
Direct Submission
Submitted (30-Mar-1998) Whitehead Institute for Biomedical
Research, Nine Cambridge Center, Cambridge, MA 02142, USA
LOCATION/Qualifiers
1. 2358
/organism="Oryza sativa"
/mol_type="mRNA"
/db_xref="taxon:4530"
1. 2358
/gene="REH1"
152. .1939
/gene="REH1"
/feature="potential membrane protein"
/codon_start=1
/product="auxin transport protein REH1"
/db_xref="GI:3377509"
/translation="MITAADFHYHVTAMVPLVYVAMLYAGSVKWRIFTDPQCSGINSR
FVALPVLPLSFHFSTNNPVTMLRFLAADTLQKLVMLAMLTAMSLSRGSLRWFI
TLFSLSTPNTLVMSIGIPLLKMGVYFSGSLMVQVLCQI IWTLMFLMPEYRGAAML
ITEQPDPTANAI IAVNDPVSVDGREDIAETETEVKEDGR IHTVTVRMSNARSDIY
SRRSGPSTTPRPNTLNAEYLSLOSSENPTRGSPNHTDFYSWVGRSNFGAADA
FGVRTGATPRSNEDDASKPKIPLIPASNRAPMAGHPAPNPVAPSNAPKKNKKAATNG
QAKGDELHMFVSSASPSVDFGGAPDYNDAAVKSPKQDKAKQREDEKDDPS
FCNRMVDRDAEAGDEKAAAGADPSKAAAPAMPPTSVMTLILIMVWRKLINP
NTYSSLI GLIWSLVCFRNFWPAIVLKSISILSDAGLGMAMFSLGLFMALQPHI IAC
GNKVATYAMVRFVLAGPAMAAAGFAVGLRGLTLLHVAIVQAALPQGI VPFVFAKEYSV
HPSILSTAVIFGMLIALPILVYIILGL"
FEATURES
source
gene
CDS

```

```

ORIGIN
Alignment Scores:
Pred. No.: 5,43e-85 Length: 2358
Score: 1665.00 Matches: 365
Percent Similarity: 67.58% Conservative: 54
Best Local Similarity: 58.87% Mismatches: 111
Query Match: 57.23% Indels: 90
DB: 8 Gaps: 17
US-10-030-884-14 (1-573) x AF056027 (1-2358)
Qy 1 MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValProLeuTyrValAla 20
Db 152 ATGATTACGGCGCGGAGCTTCTTACACGATGATGACGGCGATGCTGCTGCTTACCTTGTACCTGGCG 211
Qy 21 MetThrLeuAlaTyrGlySerValArgTrpTrpArgIlePheThrProAspGlnCysSer 40
Db 212 ATGATACCTGGCTACGGGTGCGTGAAGTGGTGGCGCATCTTTCACGCCGAGACCACTGCTCC 271
Qy 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60
Db 272 GGGATCAACCGCTTCTGTCGCGCTCTTCGCGCGCTGCTGCTGCTTCCACTTCACTTCACTCC 331
Qy 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80
Db 332 ACCACACACCGCTACACGATGACCTTCGGTTCATCGCGCGCCGACACGCTGCAGAGCTG 391
Qy 81 AlaValLeuAlaLeuLeuAla-----LeuAlaSerArgGlyLeuSerSerPro 96
Db 392 ATGGTGTCTGGCCATGCTCAGCGCGTGGAGCCACCTCAGCGCGCGGGG----- 439
Qy 97 ArgAlaLeuGlyLeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThr 116
Db 440 -----AGCTTCAGTGGACCATCAGCTCTTCTCCCTCTCCACGCTGCCACACAGC 490
Qy 117 LeuValMetGlyIleProLeuLeuArgGlyMetTyrGlyIleAlaSerSerAlaGlyThrLeu 136
Db 491 CTCGTATCGGATCCCTTTGCTCAAGGCGATGTACGGGAGTTCTTCC---GGCAGCCTC 547
Qy 137 MetValGlnValValValLeuGlnCysIleIleTyrThrLeuMetLeuPheLeuPhe 156
Db 548 ATGGTGCAGATCGTCTGCTGTCAGTGCATCATCTGTACACGCTCATGCTCTTCTATGTC 607
Qy 157 GluTyrArgAlaAlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSer 176
Db 608 GAGTACCGCGCGCGCGGATGCTCATCCGAGCAGTTCCTCCGAC---ACCAGCGCCAC 664
Qy 177 IleValSerPheArgValAspSerAspValValSerLeu-----AlaArgGlyAspVal 194
Db 665 ATCGCCTCCATCGTCTGTCGACCGGACGTCGTGTCGTGCGAGCGGAGGAGGAGCGCCATC 724
Qy 195 GluLeuGluAlaGluProAspGlyValAlaAlaGlyAlaGlyAlaValSerSerArgGlyGly 214
Db 725 GAGACGAGACGAGGAGTGAAG----- 745
Qy 215 AspAlaGlyArgValArgValThrValArgLysSerThrSerSerArgSerGluAlaAla 234
Db 746 GAGGACGCGCAGGATACACGTCGTCGCGCGCTCCACGCGCTCTCGCTCGGACATCTAC 805
Qy 235 CysSerHisSerHisSer---GlnThrMetGlnProArgValSerAsnLeuSerGlyVal 253
Db 806 TCCCGCGCTCCATGGGCTTCTCCAGCACACGCGCGCGCGCGGACACCTCCACCAACGCC 865
Qy 254 GluIleTyrSerLeuGlnSerSerArgAsnProThrProArgGlySerSerPheAsnHis 273
Db 866 GAGATCTACTCTGCTGCTGCTGCGGAAACCGGACCGCGGAGGGGTTCAGAGCTTCAACCCAC 925
Qy 274 AlaAspPhePheAsnIleValGlyAlaAlaAlaLysGlyGlyGlyAlaAlaGlyAsp 293
Db 926 ACCGATTTCTACTTCCATGTGGCGCCGACGCTCCAACTTCGCGCGCGCGCGCGCTTCGCG 985
Qy 294 GluGlyLeuGlyAlaCys----- 299

```

```

Db 986 GTCCGACCGCGCCACGCGCGCGCCCGTCCAACTACGAGGACGACGCGTCCAAAGCCCAAG 1045
QY 300 -----GlyGlyGlyGlyGlyHis-----SerProGlnPro 310
Db 1046 TACCCGCTCCCGCGTCGAAATCGCGCGCCATCGCGGCCCATCTACCGCGCGCGCAACCG 1105
QY 311 GlnAlaValAlaValPro-----AlaIysArg 319
Db 1106 GCCGTGTCTCGCGCGCCCAAGGCGCCAGAGCGCGCCACGAAAGCGCGCGCGCAAGGCG 1165
QY 320 LysAspLeuHisMetLeuValTrpSerSerAlaSerProValSerGluArgAlaAla 339
Db 1166 GAGGACCTCCCATGTTGCTGTGGAGCTCCAGCGGTCGCGCGTGTCCGAC----- 1216
QY 340 ValHisValPheGlyAlaGlyGlyAlaAspHisAlaAspValLeuAla----- 355
Db 1217 -----GTCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1270
QY 356 -----LysGlyValAlaGlnAlaValTrpSerSerAlaSerProValSerGluArgAlaAla 371
Db 1271 CCAAAATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1330
QY 372 ArgThrLysAsnGlySerGlyGlyAlaAspLysGlyGlyProThrLeuSerLysLeuGly 391
Db 1331 GCGAACAGG-----GCGTCATGACAGGAC-----GCGGAGCGCAGG 1369
QY 392 SerAsnSerThrAlaGlnLeuTyProLysAspAspGlyGluGlyArgAlaAlaVal 411
Db 1370 GACGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1429
QY 412 AlaMetProProAlaSerValMetThrArgLeuLeuLeuLeuMetValTrpArgLysLeu 431
Db 1430 GCGATGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1489
QY 432 IleArgAsnProAsnThrTyrsSerSerLeuIleGlyValValTrpSerLeuValSerTyr 451
Db 1490 ATCCGCAACCGCAACCACTTACTTCCAGCTTCTGCGCTCATCTGCTGCTGCTGCTGCT 1549
QY 452 ArgTTPGlyIleLeuMetProAlaIleLeuAlaArgSerIleSerIleLeuSerAspAla 471
Db 1550 AGTGGAACTTCGAGTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1609
QY 472 GlyLeuGlyMetAlaMetPheSerLeuGlyLeuPheMetAlaLeuGlnProArgIleIle 491
Db 1610 GGGCTCGGATGGCCATGTTTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1669
QY 492 AlaCysGlyAsnLysLeuAlaIleAlaMetGlyValArgPheValAlaGlyProAla 511
Db 1670 GCGTGGGAAACAGGTGGCGAGCGTACGCGTGGCGGTCGCGTTCCTGCGCGCGCGCG 1729
QY 512 ValMetAlaAlaAlaSerIleAlaValGlyLeuArgGlyValLeuLeuHisIleAlaIle 531
Db 1730 GTGATGGCGCGCGGTCCTTCGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGG 1789
QY 532 ValGlnAlaAlaLeuProGlnGlyIleValProPheValPheAlaLysGluTyrGlyVal 551
Db 1790 GTCCAGCGAGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1849
QY 552 HisProAspIleLeuSerThrAla-----TyrGly-----ProIleThr 564
Db 1850 CACCTAGCATTTCTAGCAGCGTGTCTTGGCATGCTCATCGCTTGGCTTATCACC 1909

RESULT 9
AK102343
LOCUS
DEFINITION
Oriza sativa (japonica cultivar-group) cDNA clone:J033091B16, full
insert sequence.
ACCESSION
AK102343
VERSION
JUI_CDNA; 1 GI:32987552
KEYWORDS
Oriza sativa (japonica cultivar-group)
CAP trapper.
ORGANISM
Oriza sativa (japonica cultivar-group)

```

REFERENCE  
AUTHORS

1 The Rice Full-length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-length cDNA Project Team: Kikuchi S., Satoh K., Nagata T., Kawagashira N., Doi K., Kishimoto N., Yazaki J., Ishikawa M., Yamada H., Hotta I., Kojima K., Namiki T., Ohneda E., Yahagi W., Suzuki K., Li C., Ohtsuki K., Shishiki T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group: Ohtsuki Y., Iida Y., Sugano S., Fujimura T., Suzuki Y., Tsunoda Y., Kuroaki T., Kodama T., Masuda H., Kobayashi M., Xie Q., Lu M., Nariawa R., Sugiyama A., Mizuno K., Yokomizo S., Niikura J., Ikeda R., Ishibiki J., Kawamata M., Yoshimura A., Miura J., Kusumegi T., Oka M., Ryu R., Ueda M., Matsubara K., RIKEN, Hara A., Carninci P., Adachi J., Aizawa K., Arakawa T., Fukuda S., Kawaguchi I., Kondo S., Konno H., Miyazaki A., Ohtsuki T., Itoh M., Saito R., Sasaki D., Sato K., Shibata K., Shinagawa A., Shiraki T., Yoshino M., and Hayashizaki Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice

Science 301 (5631), 376-379 (2003)

22752273

12869764

REFERENCE  
AUTHORS

2 (bases 1 to 2402)  
Adachi J., Aizawa K., Akimura T., Arakawa T., Carninci P., Doi K., Fujimura T., Fukuda S., Hanagaki T., Hara A., Hashizume W., Hayashida K., Hayashizaki Y., Hayatsu N., Hiramoto K., Hiraoka T., Hori F., Hotta I., Iida Y., Ikeda R., Inamura K., Iotani K., Ishibiki J., Ishii Y., Ishikawa M., Itoh M., Kagawa I., Kawaguchi S., Kato H., Kawagashira N., Kawai J., Kawamata M., Kikuchi S., Kishikawa Hirozane T., Kishimoto N., Kobayashi M., Kodama T., Kojima K., Kojima Y., Kondo S., Konno H., Kouda M., Koya S., Kurihara C., Kurosaki T., Kusumegi T., Li C., Lu M., Masuda H., Matsubara K., Matsuyama T., Miura J., Miyazaki A., Mizuno K., Murakami K., Niikura J., Nagata T., Nakamura M., Namiki T., Nariawa R., Niikura J., Nishi K., Nomura K., Numasaki R., Ohneda E., Ohno M., Ohtsuki K., Oka M., Ooka H., Oota N., Ota Y., Ohtsuki K., Ohtsuki K., Sakai C., Sakai K., Sakazume N., Sano H., Sasaki D., Sato K., Satoh K., Shibata K., Shingawa A., Shiraki T., Shishiki T., Sogabe Y., Sugano S., Sugiyama A., Suzuki K., Suzuki Y., Tagami M., Tagami-Takeda Y., Tagawa A., Takahashi F., Takaku-Akahira S., Tanaka T., Tomaru A., Toya T., Tsunoda Y., Ueda M., Waki K., Xie Q., Yahagi W., Yamada H., Yamamoto M., Yasunishi A., Yazaki J., Yokomizo S. and Yoshimura A.

Direct Submission

Submitted (27-RUG-2002) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)

This clone is one of the 28K full-length cDNA clones from japonica rice.

URL: <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-length cDNA Project Team: Kikuchi S., Satoh K., Nagata T., Kawagashira N., Doi K., Kishimoto N., Yazaki J., Ishikawa M., Yamada H., Hotta I., Kojima K., Namiki T., Ohneda E., Yahagi W., Suzuki K., Li C., Ohtsuki K., Shishiki T. and Yamamoto M.

FAIS Genome Sequencing & Analysis Group: Ohtsuki Y., Iida Y., Fujimura T., Ikeda R., Ishibiki J., Kawamata M., Kobayashi M., Kodama T., Kurosaki T., Kusumegi T., Lu M., Masuda H., Miura J., Mizuno K., Nariawa R., Niikura J., Oka M., Ryu R., Sugano S., Sugiyama A., Suzuki Y., Tsunoda Y., Ueda M., Xie Q., Yokomizo S., Yoshimura A., Matsubara K. and Murakami K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi J., Aizawa K., Akimura T., Arakawa T., Carninci P., Fukuda S., Hanagaki T., Hara A., Hashizume W., Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hori F., Iida J., Inamura K., Imotani K., Ishii Y.,

TITLE  
JOURNAL

## COMMENT

Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,  
 Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Korno, H., Kouda, M.,  
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
 Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N.,  
 Ota, Y., Saichoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,  
 Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
 Sogabe, J., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,  
 Takaku-Akashira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,  
 Yasunishi, A. and Hayashizaki, Y.  
 Location/Qualifiers  
 source  
 1..2402  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /clone="J033091B16"

## ORIGIN

## Alignment Scores:

Pred. No.: 5,56e-85 Length: 2402  
 Score: 1665.00 Matches: 365  
 Percent Similarity: 67.58% Conservative: 54  
 Best Local Similarity: 58.87% Mismatches: 111  
 Query Match: 57.22% Indels: 90  
 DB: 8 Gaps: 17

US-10-030-884-14 (1-573) x AK102343 (1-2402)

Qy 1 MetThrAlaLeuAspLeuTyrHisValLeuThrAlaValProLeuTyrValAla 20  
 Db 151 ATGATACGGCGCGGAGCTTCTTACCACTGATGACGGCGATGTCGCGTTCAGTCGCG 210  
 Qy 21 MetThrLeuAlaTyrGlySerValArgTrpTrpArgIlePheThrProAspGlnCysSer 40  
 Db 211 ATGATACGGCGGTCAGCGGTGCGTGAAGTGGTGGCGCATCTTCACCGCGGACGATGTC 270  
 Qy 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
 Db 271 GGGATCAACCGCTTCGTGGCGCTCTTCGCGGTGCGCTGCTGCTTCCATCTCATCTCC 330  
 Qy 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAspThrLeuGlnIleVal 80  
 Db 331 ACCAACACCCGTCATCAGCATGAACCTCCGCTGTCATCGCGCGGACACGCTCGAAGCTG 390  
 Qy 81 AlaValLeuAlaLeuAla-----LeuAlaSerArgGlyLeuSerSerPro 96  
 Db 391 ATGGTGTGGCCATCTCACGGCTGGAGCCACCTCAGCCCGCGGGG-----438  
 Qy 97 ArgAlaLeuGlyLeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThr 116  
 Db 439 -----AGCTTCGAGTGGACCATCAGCTCTTCTCCCTCTCCACGCTGCCCAACACG 489  
 Qy 117 LeuValMetGlyLeuProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeu 136  
 Db 490 CTCGTCATGGGATCCCTTGTCTAAGGCGATGTCAGGGAGTCTCC---GGCAGCTC 546  
 Qy 137 MetValGlnValValLeuGlnCysIleIleTrpTyrThrLeuMetLeuPheLeuPhe 156  
 Db 547 ATGGTGCAGATCGTGTGTCGAGTGTCATCTGCTACAGCTCATGCTCTTCTTCATGTC 606  
 Qy 157 GluTyrArgAlaAlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSer 176  
 Db 607 GAGTACCGCGGCGCGCGATGTCATACCGAGCATGTCCTCCGAGC-----ACCGCGCCCAAC 663  
 Qy 177 IleValSerPheArgValAspSerAspValValSerLeu-----AlaArgGlyAspVal 194  
 Db 664 ATCGCTCCATCGTGTGACCGCGGACGTCGTGTCGCTGGACGGCAGGAGGAGCGCCATC 723  
 Qy 195 GluLeuGluAlaGluProAspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGly 214  
 Db 724 GAGACGGAGACGAGGTGAAG-----744  
 Qy 215 AspAlaGlyArgValArgValThrValArgLysSerThrSerSerArgSerGluAlaAla 234

Db 745 GAGGACGGGAGGATACACGTCACCGTGGCCCGCTCCAAACGCGTCTGCTCGGACATCTAC 804  
 Qy 235 CysSerHisSerHisSer---GlnThrMetGlnProArgValSerAsnLeuSerGlyVal 253  
 Db 805 TCCCGCGCTCCATGGCTTCTCCAGCACCAACGCGCGCGGCGCAACCTCACCACCGCC 864  
 Qy 254 GluIleTyrSerLeuGlnSerSerArgAsnProThrProArgGlySerSerPheAsnHis 273  
 Db 865 GAGATCTACTCGTCTGTCAGTCGTCGCGAAACCGGACCGCGGAGGGTTCAAGCTTCAACCAC 924  
 Qy 274 AlaAspPhePheAsnIleValGlyAlaAlaAlaLysGlyGlyGlyAlaAlaGlyAsp 293  
 Db 925 ACCGACTTCTACTCCATGTTGGCGCGCAGCTCCAACTTCGGCGCGCCGACGCGTTCGGC 984  
 Qy 294 GluGluLysGlyAlaCys-----299  
 Db 985 GTCCGACCGCGGCCACGCGCGCGCTCCAACTACGAGGACGACGCTCCAAAGCCCAAG 1044  
 Qy 300 -----GlyGlyGlyGlyGlyHis-----SerProGlnPro 310  
 Db 1045 TACCGCTCCGGCGTCGAATGGCGGCCCATGGCGGCATCTACCGCGCGCGACCCG 1104  
 Qy 311 GlnAlaValAlaValPro-----AlaLysArg 319  
 Db 1105 GCGGTCTGTCGCGCGCCCAAGGCGGCCAACGAGCGGCACGAAACGCGGAGGCTCAAGGGC 1164  
 Qy 320 LysAspLeuHisMetLeuValTrpSerSerSerAlaSerProValSerGluArgAlaAla 339  
 Db 1165 GAGGACCTCCACATGTCGTGAGCTCCAGCGCTCGCGCTCGCTGTCGCG-----1215  
 Qy 340 ValHisValPheGlyAlaGlyAlaAspHisAlaAspValLeuAla-----355  
 Db 1216 -----GTCTTGGCGCGCGCGCCAGACTCAACAGCGCGCGCGCTCAAGTCCCCC 1269  
 Qy 356 -----LysGlyAlaGlnAlaTyrAspGluTyr---GlyArgAspAspTyrSerSer 371  
 Db 1270 CGCAAAATCGATGGACGGAAGGACGAGGAGTACGTGGAGCGGGACCATTTTCAGCTTC 1329  
 Qy 372 ArgThrLysAsnGlySerGlyAlaAspLysGlyGlyProThrLeuSerLysLeuGly 391  
 Db 1330 GGGAAACAGG-----GGCGTCATGGACAGGAC-----CGGAGGCGAGG 1368  
 Qy 392 SerAsnSerThrAlaGlnLeuTyrProLysAspAspGlyGlyArgAlaAlaAlaVal 411  
 Db 1369 GACGAGAGCG 1428  
 Qy 412 AlaMetProProAlaSerValMetThrArgLeuLeuLeuMetValTrpArgLysLeu 431  
 Db 1429 GCGATCCGCGCGACGAGCGGTGATGACCCGCTCATCTCTGATCATGTGTGGCGCAAGCTC 1488  
 Qy 432 IleArgAsnProAsnThrTyrSerSerLeuIleGlyValValTrpSerLeuValSerTyr 451  
 Db 1489 ATCCGCAACCGCAACCTTACTCCAGCTCATCGGCTCATCTGCTCCCTCGTCTCTCTTC 1548  
 Qy 452 ArgTrpGlyIleGluMetProAlaIleAlaArgSerIleSerIleLeuSerAspAla 471  
 Db 1549 AGGTGAACATTCAGATGCGCGCATCTGCTCTGTAATCATCTCTGATCTCTGTCGCGAGCG 1608  
 Qy 472 GlyLeuGlyMetAlaMetPheSerLeuGlyLeuPheMetAlaLeuGlnProArgIleIle 491  
 Db 1609 GGGCTGGCATGGCCATGTTCACTCTCGTCTGTTTCATGGCTGCAGCGGCACATCATC 1668  
 Qy 492 AlaCysGlyAsnLysLeuAlaAlaIleAlaMetGlyValArgPheValAlaGlyProAla 511  
 Db 1669 GCGTGGCGGAACAAAGTGGCGCATGCGCATGGCGTGGGTCTCTGGCGCGCGCGCGCC 1728  
 Qy 512 ValMetAlaAlaAlaSerIleAlaValGlyLeuArgGlyValLeuLeuHisIleAlaIle 531  
 Db 1729 GTGATGGCGCGCGCTCTCTCGCGCTCGGACTCCGTGGCGACGCTCTCTGACGCTCGCAT 1788  
 Qy 532 ValGlnAlaAlaLeuProGlnGlyIleValProPheValPheAlaLysGluTyrGlyVal 551



Db 1789 GTCCAGGACGCTTCTCCCGCAGGGGCAATGTGCTCCCTTGTGCTTCCCGAAGAGGTACAGCGTGT 1848

Qy 552 HisProAspIleLeuSerThrAla-----Tyr-Gly-----ProIleThr 564  
|||||  
Db 1849 CACCCTTAGCATTTCTCAGCACAGCTGTCTATCTTTGGCATGCTCATCGCTTCGCTATCACC 1908

RESULT 10  
AK101504  
LOCUS Oryza sativa (japonica cultivar-group) cDNA clone:J033044E23, full  
DEFINITION insert sequence.  
AK101504  
ACCESSION AK101504.1 GI:32986713  
VERSION FLI CDNA; CAP t-rapper.  
KEYWORDS Oryza sativa (japonica cultivar-group)  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Eriaristaceae; Oryzaceae; Oryza.  
1

REFERENCE  
AUTHORS The Rice Full-Length cDNA Consortium, National Institute of  
Agrobiological Sciences Rice Full-Length cDNA Project Team,  
Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,  
Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,  
Kojima, K., Naniiki, T., Ohneda, E., Yanagi, W., Suzuki, K., Li, C.,  
Ohtsuki, K., Shishiki, T., Foundation of Advancement of International  
Science Genome Sequencing & Analysis Group; Otsu, Y., Murakami, K.,  
Lida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,  
Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,  
Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Nishikura, J.,  
Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,  
Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN,  
Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,  
Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,  
Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y.,  
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
Yoshino, M., and Hayashizaki, Y.  
Collection, mapping, and annotation of over 28,000 cDNA clones from  
japonica rice  
Science 301 (5631), 376-379 (2003)  
27525273  
22869764  
2 (bases 1 to 2457)

REFERENCE  
AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,  
Fujimura, T., Fukuda, S., Hangaki, T., Hara, A., Hashizume, W.,  
Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramatsu, K., Hirata, T.,  
Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,  
Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,  
Kawagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,  
Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,  
Kodama, I., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,  
Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,  
Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,  
Naniiki, T., Narikawa, R., Nishikura, J., Nishi, K., Nomura, K.,  
Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,  
Osato, N., Ota, Y., Otonari, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,  
Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,  
Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,  
Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,  
Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,  
Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yanagi, W.,  
Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and  
Yoshimura, A.  
Direct Submission

TITLE Submitted (27-AUG-2002) Shoichi Kikuchi, National Institute of  
JOURNAL Agrobiological Sciences, Department of Molecular Genetics, Head of  
Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki  
305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp)  
Tel:81-29-838-7007, Fax:81-29-838-7007  
This clone is one of the 28K full-length cDNA clones from japonica  
rice.  
URL : <http://cdna01.dna.affrc.go.jp/cDNA/>



```

769 GTCTGTCCTCCAGTGCATCATCTGGTACACGCTGATGCTCTTCCACATTCGAGTTCGCGGCC 828
161 AlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSerIleValSerPhe 180
829 GCGCGGATGCTGATCGCGGACGACAGTTTCGGAC---ACGGGGGGTCCATCGTGTCTCCCTG 885
181 ArgValAspSerAspValValSerLeuAlaArgGlyAspValGluLeuGluAlaGluPro 200
886 CACGTGCAACCGCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 942
201 AspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGlyAspAlaGlyArgValArg 220
943 -----GTGGCGGG-----GACGGCGGCTGCAC 966
221 ValThrValArgLysSerThrSerArgSerGluAlaAlaCysSerHisSerHisSer 240
967 GTACCGTGGCGCGGCTCTCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1002
241 GlnThrMetGlnProArgValSerAsnLeuSerGlyValGluIleTyrSerLeuGlnSer 260
1003 CTGCTGGTCAACCGCGCGGCTCGAACCTGACGGAGCGGAGATCTACTCTCTTAGCTTCG 1062
261 SerArgAsnProThrProArgGlySerSerPheAsnHisAlaAspPheAsnIleVal 280
1063 TCGCGGAACCCCAACCCCGCGGCTCCCACTTCAACCAACCGCGACTTCTTCCGCAATGTC 1122
281 Gly-----AlaAlaAlaLysGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 292
1123 GCGCGCGGCGCCACCGCCCGCCCGCGCGCTCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1182
293 Asp-----GluGlu 295
1183 GAGCTCTACTCGCTGCAATCGTGGCGGGGCCCAACCGGAGGAGTCCCACTTCGACGAG 1242
296 LysGlyAlaCysGlyGlyGlyGlyGlyGlyHisSerProGlnProGlnAlaValAlaVal 315
1243 CACTCGGCA-----CGGCCACCGAAGAACCCACCGCGGACGACCCACG 1281
316 ProAla-----LysArgLysAspLeuHisMetLeuValTyrSerSerSerAlaSer 332
1282 GGGGCACTCAACGAGTGCAGGAGTCCCAATGTTCTGTTGAGTGTGAGTGTGAGGCGGTAT 1341
333 ProValSerGluArgAlaAlaValHisValPhe-----GlyAlaGlyGlyAlaAspHisAla 351
1342 CCGGTCTCAGAGTCAAGCGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1401
352 AspValLeuAlaLys----- 356
1402 GACGTGGCGCGCAAGAAATCCACATGTTATCCCGCGGACCTCCCGCAGAACACCGC 1461
357 GlyAlaGlnAlaTyrAspGluTyrGlyArgAspAspTyrSerSerArgThrLysAsnGly 376
1462 TCAGGCAAGACGACGAGGAGTACGGC-----GCAGTGGCATTTGGGTGGC 1506
377 SerGlyGlyAlaAspLys-----GlyGlyProThr----- 386
1507 GCGCGCGCGGAGAGAACTTCAGTTTCGGAGGCGGCAAGAGCGTGGCGCGCGCGAGGCA 1566
387 -----LeuSerLysLeuGlySerAsnSerThrAla 396
1567 GTAGACGAGAGCGGCGCTTGTGCTGACGGCTGACGAGATGGGTTCGAGTTCGACGCGG 1626
397 GlnLeuTyrProLys-----AspAspGlyGluGlyArgAlaAla 409
1627 GAGTGCACCCGAGGTCGTGACGTCGACGCGGACCGAGCGCGCGCGCGCGCGCGCGCGCG 1686
410 Ala-----ValAlaMetProAlaSerValMetThrArgLeuIleLeuIleMetVal 427
1687 GCGGGGCGAGTACCAATGCGCGCGCGGCGGAGTATGACAGCGCTCATCTCATATATGTTG 1746
428 TrpArgLysLeuIleArgAsnProAsnThrTyrSerSerLeuIleGlyValValTrpSer 447

```

```

1747 TGGCGGCAAGCTCATTCGGCAACCCCAACACATTACTCAGCGCTCTCGGCTCGCTGGTCC 1806
448 LeuValSerTyrArgTrpGlyIleGluMetProAlaIleAlaAlaArgSerIleSerIle 467
1807 CTGCTCGCTTCGGTGGCAGCGTCTCCATCGCAGCAATCGTCGAGAAAGTCCATCTCCATT 1866
468 LeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeuGlyLeuPheMetAlaLeuGln 487
1867 CTCTCGAGCGAGCGCTGGGATGGCCATGTTTAGCTGGGATGTTTCTATGGCGCTGCAG 1926
488 ProArgIleIleAlaCysGlyAsnLysLeuAlaAlaIleAlaMetGlyValArgPheVal 507
1927 CCCAGCATCATCGGTGTGGCAATCAGCGCGCGTCTCCATCGCGCTCGCTTCCTC 1986
508 AlaGlyProAlaValMetAlaAlaSerIleAlaValGlyLeuArgGlyValLeuLeu 527
1987 GCGGGCGCTCGCGTATGCGCGCGGTCAATCGCATCGGACTCCGCGGAGCGCTCTCTG 2046
528 HisIleAlaIleValGlnAlaAlaLeuProGlnGlyIleValProPheValPheAlaLys 547
2047 CAGTCCGCAATGTTTCAGCGCGCTCTACCAAGGATGTGCTTTGTTTGGCAAA 2106
548 GluTyrGlyValHisProAspIleLeuSerThrAla-----TyrGly----- 561
2107 GAATACAATGTCCACCGCGCATCTCTGAGCACAGCGGTAAATTTTGGCATGCTAATAGCT 2166
562 ---ProIleThr 564
2167 CTTCCAATCA 2178

RESULT 11
AK099634      2343 bp mRNA linear PLN 24-JUL-2003
LOCUS        Oryza sativa (japonica cultivar-group) cDNA clone:J013059P15, full
DEFINITION   insert sequence.
ACCESSION    AK099634.1 GI:32984843
VERSION      AK099634.1
KEYWORDS     FLI_CDNA; CAP trapper.
SOURCE       Oryza sativa (japonica cultivar-group)
ORGANISM     Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.

REFERENCE
AUTHORS      The Rice Full-length cDNA Consortium, National Institute of
              Agrobiological Sciences Rice Full-length cDNA Project Team;
              Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,
              Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,
              Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,
              Ohtsuki,K., Shishiki,T., Foundation of Advancement of International
              Science Genome Sequencing & Analysis Group: Ohtomo,Y., Murakami,K.,
              Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,
              Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,
              Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J.,
              Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,
              Kusunegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN;
              Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,
              Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Tshii,Y., Itoh,M.,
              Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y.,
              Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
              Yoshino,M. and Hayashizaki,Y.
              Collection, mapping, and annotation of over 28,000 cDNA clones from
              japonica rice
              Science 301 (5631), 376-379 (2003)

TITLE
JOURNAL      Science 301 (5631), 376-379 (2003)
MEDLINE      22752273
PUBMED       12869764
REFERENCE    2 (bases 1 to 2343)
AUTHORS      Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,
              Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,
              Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T.,
              Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K.,
              Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,
              Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M.,

```

COMMENT

Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Negata, I., Nakamura, M., Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, Y., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Otsu, Y., Ryo, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Teunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

Direct Submission

Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression, 2-1-2 Kamondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel: 81-29-838-7007, Fax: 81-29-838-7007)

This clone is one of the 28K full-length cDNA clones from japonica rice.

URL: <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Sato, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Fotta, I., Kojima, K., Namiki, T., Ohneda, Y., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., and Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Otsu, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryo, R., Sugano, S., Sugiyama, A., Suzuki, Y., Teunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, E., Iida, J., Imamura, K., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Otsu, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

Location/Qualifiers

FEATURES

source  
1. .2343  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="J013059F15"

ORIGIN

Alignment Scores:  
Pred. No.: 1,73e-84 Length: 2343  
Score: 1636.00 Matches: 364  
Percent Similarity: 67.42% Conservative: 54  
Best Local Similarity: 58.71% Mismatches: 112  
Query Match: 56.91% Indels: 90  
DB: 8 Gaps: 17

US-10-030-884-14 (1-573) x AK099634 (1-2343)

Qy 1 MetThrAlaLeuAspLeuTyrHisValLeuThrAlaValValProLeuTyrValAla 20  
Db 151 ATGATTACGGCGGGAGCTTCTACACGTGACGGCGATGGTGGCGTTGTACGGGG 210  
Qy 21 MetThrLeuAlaTyrGlySerValArgTyrTrpArgIlePheThrProAspGlnCysSer 40

Db 211 ATGATACTGGCGTACGGGTCGGTAAAGTGGTGGCGCATCTTTCACGGCGGACCAAGTCTCC 270  
Qy 41 GlyTleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
Db 271 GGGATCAACCGCTTCGTGGGCTCTTCGGCGTGGCGCTGCTGCTGCTTCCACTTCACTCC 330  
Qy 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnVal 80  
Db 331 ACCAACACCGGTACACGATGAACCTCCGGTTCATCGCGCGACACGCTGCAAGAGCT 390  
Qy 81 AlaValLeuAlaLeuLeuAla-----LeuAlaSerArgGlyLeuSerSerPro 96  
Db 391 ATGGTGTGGCCATGCTCAGGGGTGGAGCCACTTCAGCGCGGGGG----- 438  
Qy 97 ArgAlaLeuGlyLeuAspTrpSerIleThrLeuPheSerIleuSerThrLeuProAsnThr 116  
Db 439 -----AGCCTCGAGTGGACCATCAGCTCTTCTCCCTCTCCACGCTGCCACACG 489  
Qy 117 LeuValMetGlyTleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeu 136  
Db 490 CTGTCTATGGGATCCCTTGTCAAGGCGATGTACGGGAGTTCTCC---GGCAGCCTC 546  
Qy 137 MetValGlnValValValLeuGlnCysIleIleTyrTrpThrLeuMetLeuPheLeuPhe 156  
Db 547 ATGGTGCAGATCGTGTGCTGAGTGCATCATCTGTTACACGCTCATGCTCTTTCATGTT 606  
Qy 157 GluTyrArgAlaAlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSer 176  
Db 607 GAGTACCGCGCGCGCGATGCTCATCCGCGAGTTCCCGAC---ACCGCGCGCAAC 663  
Qy 177 IleValSerPheArgValAspSerAspValValSerIleu-----AlaArgGlyAspVal 194  
Db 664 ATGCCTCCATCGTGTGCGACCGCGAGCTGTGTGCGACGCGAGGAGGACGCATC 723  
Qy 195 GluLeuGluAlaGluProAspGlyValAlaGlyAlaValSerSerArgGlyGly 214  
Db 724 GAGACGAGAGCGAGGTGAAG----- 744  
Qy 215 AspAlaGlyArgValArgValThrValArgValSerThrSerArgSerGluAlaAla 234  
Db 745 GAGGACGCGAGGATACACGTCACGTCGCGCGCTCCAAACGCTCTCGCTCGCATCTAC 804  
Qy 235 CysSerHisSerHisSer---GlnThrMetGlnProArgValSerIleuSerArgVal 253  
Db 805 TCCCGCGCGCTCCATGGGCTTCTCCAGCACCGCGCGCGCGCGAGCACTTCCACCAAGCC 864  
Qy 254 GluIleTyrSerLeuGlnSerSerArgAsnProThrProArgGlySerSerPheAsnHis 273  
Db 865 GAGATCTACTCGTGTGAGTGTGCGGACCGCGAGCGCGGAGGGTTCAAGCTTCAACCCAC 924  
Qy 274 AlaAspPhePheAsnIleValGlyAlaAlaAlaValGlyGlyGlyAlaAlaGlyAsp 293  
Db 925 ACCGACTTCTACTCGTGGCGCGAGCTTCCAACTTCGCGCGCGCGCGCGCTTCGCGC 984  
Qy 294 GluGlyGlyGlyAlaCys----- 299  
Db 985 GTCGCGACCGCGCGCGCGCGCGCTCCAACTACGAGGACGCGCTCCCAAGCCCAAG 1044  
Qy 300 -----GlyGlyGlyGlyGlyHis-----SerProGlnPro 310  
Db 1045 TACCGCTTCCCGCGTCCGAATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1104  
Qy 311 GlnAlaValAlaValPro-----AlaIleArg 319  
Db 1105 GCGGTGTGTGCG 1164  
Qy 320 LysAspLeuHisMetLeuValTrpSerSerSerAlaSerProValSerGluArgAlaAla 339  
Db 1165 GAGGACCTCCACATGTTCTGCTGAGCTCCAGCGCTGCGCGCTGCTGCGAC----- 1215  
Qy 340 ValHisValPheGlyAlaGlyGlyAlaAspHisAlaAspValLeuAla----- 355

1216	Db	-----GTCTTCGGCGCGCGCGCCAGACTCAACAGCGCCGCGGAGTCAAGTCCCC	12169
356	Qy	-----LysGlyAlaGlnAlaTyAspGluTyr---GlyArgAspAspTyrSerSer	371
1270	Db	CGCAAAATGATGAGCGAAGGACAGGAGGACTACGTGGAGCGGACGATTTTCAGCTTC	1329
372	Qy	ArgThrLysAsnGlySerGlyGlyAlaAspGlyGlyProThrLeuSerLysLeuGly	391
1330	Db	GGGAAACAGG-----GCGGTATGACAGGGAC-----CGGAGGCGAGGG	1368
392	Qy	SerAsnSerThrAlaGlnLeuTyrProLysAspGlyGluGlyArgAlaAlaVal	411
1369	Db	GACGAGAGCGCGCGCGCGCGCGCGACCCAGCAGGCCATGGCGCGCGCAGC	1428
412	Qy	AlaMetProProLysValMetThrArgLeuLeuLeuMetValTrpArgLysLeu	431
1429	Db	GCGATGCGCGCGAGCGAGCGTGATGACCGCGCTCATCTCTGATCATGTGTGGCCAAAGCTC	1488
432	Qy	IleArgAsnProAsnThrTyrSerSerLeuIleGlyValValTrpSerLeuValSerTyr	451
1489	Db	ATCGCAACCGAACCTACTCTCAGCGCTCATCGGCTCATCTGGTCCCTCGTCTGCTTC	1548
452	Qy	ArgTrpGlyLeuGluMetProAlaIleIleAlaArgSerIleSerIleLeuSerAspAla	471
1549	Db	AGGTGGAACATTCAGATGCGCGGCATCGTCTGAAATCCATCTCGATCTCTGTCGAGCGCG	1608
472	Qy	GlyLeuGlyMetAlaMetPheSerLeuGlyLeuPheMetAlaLeuGlnProArgIleIle	491
1609	Db	GGGCTCGCATGGCCATGTTTCACTCTCGGTCTGTTTATGGCGGTGCAGCGCACATCATC	1668
492	Qy	AlaCysGlyAsnLysLeuAlaIleAlaMetGlyValArgPheValAlaGlyProAla	511
1669	Db	GCGTGCGGGAACAAGGTGGCGACGTACGCCATGGCGGTGCGGTCTCTGGCGCGCGCGGCC	1728
512	Qy	ValMetAlaAlaSerIleAlaValGlyLeuArgGlyValLeuLeuHisIleAlaIle	531
1729	Db	GTGATGCGCGCGCGTCTTTCGCGCTCGAGATCCGTGCGACGCTCTTCGACGTCGCCATT	1788
532	Qy	ValGlnAlaAlaLeuProGlnGlyIleValProPheValPheAlaLysGluTyrGlyVal	551
1789	Db	GTCAGGCGAGCTCTGCCCGGAGCATTTGCCCTTCGTCTCGCTAGGAGTACAGCGTG	1848
552	Qy	HisProAspIleLeuSerThrAla-----TyrGly-----ProIleThr	564
1849	Db	CACCTTAGCATTTTCACACAGCTGTCTATCTTTGGCATGTCATCGCGCTTGCCTATCACC	1908
RESULT_12			
AK063976	LOCUS	AK063976	
	DEFINITION	Oryza sativa (japonica cultivar-group) cDNA clone:001-124-C02, full insert sequence.	
	ACCESSION	AK063976	
	VERSION	AK063976.1 GI:32973994	
	KEYWORDS	FLI CDNA; oligo-capping.	
	SOURCE	Oryza sativa (japonica cultivar-group)	
	ORGANISM	Oryza sativa (japonica cultivar-group)	
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.	
REFERENCE		1	
AUTHORS			
			The Rice Full-length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team, Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Hotta,I., Ohtsuki,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Kishimoto,N., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group, Otono,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M., Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J., Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J., Kuemegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN, Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,

## Alignment Scores:

Pred. No.: 1.64e-84 Length: 2126  
 Score: 1655.50 Matches: 375  
 Percent Similarity: 66.93% Conservative: 42  
 Best Local Similarity: 60.19% Mismatches: 103  
 Query Match: 56.89% Indels: 103  
 DB: 8 Gaps: 17

US-10-030-884-14 (1-573) x AK063976 (1-2126)

QY 1 MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValProLeuTyrValAla 20  
 DB 108 ATGATATCCGGGACGACTTCTACACGGTGTGGGGCGGTGGTGGTGTACGTGGCG 167  
 QY 21 MetThrLeuAlaTyrGlySerValArgTTPToArgIlePheThrProAspGlnCysSer 40  
 DB 168 ATGTTCTTGGCGTACGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT 227  
 QY 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
 DB 228 GGATCAACCGCTTCGTCCGCTTCCTCGCGTGGCGTCTCTCTCTCTCTCTCTCTCT 287  
 QY 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnVal 80  
 DB 288 ACCAAGACCGGTACGCCATGACCTCCGCTTCCTGGCGGCGGACACGCTGCAGAGCTG 347  
 QY 81 AlaValIleAlaLeuAlaLeuAlaLeuAlaSerArgGlyLeuSerProArgAlaLeuGly 100  
 DB 348 CTGCTCTTGGCGGGTTCGCGCGGTGGTGGCGCTCCCTTCGCGACCGCGCGCGCGG 407  
 QY 101 LeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThrLeuValMetGly 120  
 DB 408 CTGGACTGGTCCATCATCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 467  
 QY 121 IleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeuMetValGlnVal 140  
 DB 468 ATCCCGTGTGATCGCCATGACGGCCATCTCC---GGCTCGCTCATGTTCAGATC 524  
 QY 141 ValValLeuGlnCysIleIleTyrThrLeuMetLeuPheLeuPheLeuGluTyrArgAla 160  
 DB 525 GTCTGTCTCCAGTGATCATCTGTGTACAGCTGATCTCTCTCTCTCTCTCTCTCTCT 584  
 QY 161 AlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSerIleValSerPhe 180  
 DB 585 GCAGGATGTGATCGCCGACCACTTCCCGGAC---ACGGCGCGCTCCATCGTGTCCCTG 641  
 QY 181 ArgValAspSerAspValValSerLeuAlaArgGlyAspValGluLeuGluAlaGluPro 200  
 DB 642 CACGTGACCCGGACGTGGTGTCTGTGGAGGGCGGCACGCGAGACGGAGCGGAG--- 698  
 QY 201 AspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGlyAspAlaGlyArgValArg 220  
 DB 699 -----GTGGCGCG-----GACGGCGCGCTGGCAC 722  
 QY 221 ValThrValArgLysSerThrSerSerArgSerGluAlaAlaCysSerHisSer 240  
 DB 723 GTACCGGTGGCGGTCTCTCGGTGTCCGG-----CGGTCTG 758  
 QY 241 GlnThrMetGlnProArgValSerAsnLeuSerGlyValGluIleTyrSerLeuGlnSer 260  
 DB 759 CTGTGTGTACCGCGCGCGGTCTCGAACCCTGACGGGACGGAGATCTACTCGCTTAGCTCG 818  
 QY 261 SerArgAsnProThrProArgGlySerSerPheAsnHisAlaAspPhePheAsnIleVal 280  
 DB 819 TCGGGGAACCAACCCCGGGGTCTCAACTTCAACCAACCGCCGACTTCTTGGCARGGTTC 878  
 QY 281 Gly-----AlaAlaAlaLysGlyGlyGlyAlaAlaGly 292  
 DB 879 GCTCC 938  
 QY 293 Asp-----GluGlu 295  
 DB 939 GAGCTCTACTCGCTGCAATCGTCTGGGGGCCCAACCCCGAGGCGAGTCCCACTTCGACGAG 998  
 QY 296 LysGlyAlaCysGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 315  
 DB 999 CACTCGGCA-----CGGCCCGCGAAACACCGCGGCAACGACGACG 1037  
 QY 316 ProAla-----LysArgLysAspLeuHisMetLeuValTrpSerSerSerAlaSer 332  
 DB 1038 GGGGCACTCAACCAACGATGCCAAGGAGCTCCACATGTCTGTGTGGAGTCTCGAGCGGTCT 1097  
 QY 333 ProValSerGluArgAlaAlaValHisValPheGlyAlaGlyGlyAlaAspHisAlaAsp 352  
 DB 1098 CCGCTCTCAAGATCAGCGGCTCTGCTGTGTTCAGTGGCGCGCGCGCGCGCTCTCGAC 1157  
 QY 353 ValLeuAlaLys-----Gly 357  
 DB 1158 GTGGCGCGAAGGAATCCACATGCTCCCGCGACCTCGCGGAGAACAAACGGGTCA 1217  
 QY 358 AlaGlnAlaTyrAspGluTyrGlyArgAspAspTyrSerSerArgThrLysAsnGlySer 377  
 DB 1218 GCGAAGACGACGAGGAGTACGCG-----GCAGTGGCATTTGGGTGGC 1259  
 QY 378 GlyGlyAlaAspLysGlyGlyProThrLeuSerLysLeuGlySerAsnSerThrAlaGln 397  
 DB 1260 GCGGEC-----GGCGGAGAGAACTTCAGC-----TCGACGGGCGAG 1295  
 QY 398 LeuTyrProLys-----AspAspGlyGluGlyArgAlaAlaAla 410  
 DB 1296 CTGACCGGAGGTCGTCTGACGCTCGACGCGACCGCGCGCGCGCGCGCGCGCGCG 1355  
 QY 411 -----ValAlaMetProProAlaSerValMetThrArgLeuIleLeuIleMetValTrp 428  
 DB 1356 GGGCAGTACCAATGCGCGCGCGGAGGTGTATGACAGCTCATCTCTCATATATGGTGTGG 1415  
 QY 429 ArgLysLeuIleArgAsnProAsnThrTyrSerSerLeuIleGlyValValTrpSerLeu 448  
 DB 1416 CGCAAGCTCATCCGCAACCCCAACACACTTACTTCCAGCTCTCTCGCTCGCTCGCTCT 1475  
 QY 449 ValSerTyrArgTTPGlyIleGluMetProAlaIleAlaArgSerIleSerIleLeu 468  
 DB 1476 GTCCCTTCGGTGGCGACGCTCCATGCCAGCAATCTCGAAGAGTCCATCTCCATCTC 1535  
 QY 469 SerAspAlaGlyLeuGlyMetAlaMetPheSerLeuGlyLeuPheMetAlaLeuGlnPro 488  
 DB 1536 TCGGACGCGAGCTGGGATGGCCATGTTAGCTGGGATGTTTATGCGGCTGCGACGCC 1595  
 QY 489 ArgIleIleAlaCysGlyAsnLysLeuAlaAlaIleAlaMetGlyValArgPheValAla 508  
 DB 1596 AGCATCATCGGTGGCAAAATCAGCGCGCTCTCTCATGGCCGCTCCGCTTCTCTCGG 1655  
 QY 509 GlyProAlaValMetAlaAlaAlaSerIleAlaValGlyLeuArgGlyValLeuLeuHis 528  
 DB 1656 GGCCTGCGCTCATGGCGCGCGCTCAATCGCATCGACTCCGCGGAGCGCTCTCTCGAC 1715  
 QY 529 IleAlaIleValGlnAlaAlaLeuProGlnGlyIleValProPheValPheAlaLysGlu 548  
 DB 1716 GTCCGCAATGTTTCAGCGGCTCTACCAAGGGATGTTGCTTTGTTTTGCAAAAGAA 1775  
 QY 549 TyrGlyValHisProAspIleLeuSerThrAla-----TyrGly----- 561  
 DB 1776 TACATGTCCACCGCGCATCTCTGAGCACAGCGGTAAATTTTGGCATGCTAATAGCTCTT 1835  
 QY 562 ProfileThr 564  
 DB 1836 CCAATCA 1844

## RESULT 13

AKI03208

LOCUS

DEFINITION

Oryza sativa (japonica cultivar-group) cDNA clone: J033122I23, full

insert sequence.

ACCESSION

AKI03208

VERSION

AKI03208.1 GI:32988417

AKI03208 2470 bp mRNA linear PLN 24-JUL-2003

Oryza sativa (japonica cultivar-group) cDNA clone: J033122I23, full

insert sequence.

ACCESSION AKI03208

VERSION AKI03208.1 GI:32988417

## KEYWORDS

FLI CDNA; CAP trapper.

## SOURCE

Oryza sativa (japonica cultivar-group)

## ORGANISM

Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzeae; Oryza.

## REFERENCE

## AUTHORS

1 The Rice Full-Length cDNA Consortium, National Institute of  
Agricultural Sciences Rice Full-Length cDNA Project Team,  
Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,  
Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,  
Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,  
Ohtsuki, K., Shishiki, T., Foundation of Advancement of International  
Science Genome Sequencing & Analysis Group, Ootomo, Y., Murakami, K.,  
Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,  
Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,  
Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,  
Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,  
Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN,  
Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,  
Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,  
Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Ootomo, Y., Ota, Y.,  
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
Yoshino, M., and Hayashizaki, Y.  
Collection, mapping, and annotation of over 28,000 cDNA clones from  
japonica rice  
Science 301 (5631), 376-379 (2003)  
2752273  
12869764

2 (bases 1 to 2470)  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,  
Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,  
Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,  
Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,  
Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,  
Kanagawa, S., Kato, H., Kawagashira, N., Kawai, J., Kawamata, M.,  
Kikuchi, S., Kishikawa, Hironaka, T., Kishimoto, N., Kobayashi, M.,  
Kodama, R., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Konda, M.,  
Koya, S., Kurihara, C., Kuroaki, T., Kusumegi, T., Li, C., Lu, M.,  
Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,  
Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,  
Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K.,  
Numasaki, R., Ohneda, E., Ohtsuki, K., Oka, M., Ooka, H.,  
Ootomo, N., Ota, Y., Ootomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,  
Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,  
Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,  
Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,  
Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,  
Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,  
Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S., and  
Yoshimura, A.  
Direct Submission  
Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of  
Agricultural Sciences, Department of Molecular Genetics, Head of  
Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki  
305-8602, Japan [E-mail: skikuchi@nias.affrc.go.jp,  
Tel: 81-29-838-7007, Fax: 81-29-838-7007]  
This clone is one of the 28K full-length cDNA clones from japonica  
rice.  
URL : <http://cdna01.dna.affrc.go.jp/cDNA/>  
NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K.,  
Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,  
Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,  
Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., and  
Yamamoto, M.  
FAIS Genome Sequencing & Analysis Group: Ootomo, Y., Iida, Y.,  
Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,  
Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,  
Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S.,  
Sugiyama, A., Suzuki, K., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,  
Yoshimura, A., Matsubara, K., and Murakami, K.  
Genome Exploration Research Group in Riken Genomic Sciences Center  
and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,

Kimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,  
Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,  
Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,  
Itoh, M., Kagawa, I., Kanagawa, S., Kato, H., Kawai, J., Konda, M.,  
Kishikawa-Hironaka, T., Kojima, Y., Kondo, S., Konno, H., Koya, S.,  
Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ootomo, N.,  
Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,  
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,  
Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,  
Yasunishi, A., and Hayashizaki, Y.  
Location/Qualifiers  
1. 2470  
source  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="J033122123"  
ORIGIN  
Alignment Scores: Length: 2470  
Pred. No.: 1,97e-84 Matches: 377  
Score: 1655.50 Conserv: 47  
Percent Similarity: 68.28% Mismatches: 103  
Best Local Similarity: 60.71% Indels: 95  
Query Match: 56.89% Gaps: 20  
DB:  
US-10-030-884-14 (1-573) x AK103208 (1-2470)  
QY 1 MetIIThrAlaLeuAspLeuThrHisValLeuThrAlaValValProLeuThrValAla 20  
DB 139 ATGATTAACGGGGCGGACATTCACCGATGATGATGATGATGATGATGATGATGATG 198  
QY 21 MetThrLeuAlaTyrGlySerValArgTptTptArgIlePheThrProAspGlnCysSer 40  
DB 199 ATGATCTGGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 258  
QY 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheLeuSer 60  
DB 259 GGGATCAACGGTTCGTGGCGTCTTTCGGCGTCTTTCGGCGTCTTTCGGCGTCTTTC 318  
QY 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnIleVal 80  
DB 319 ACCAACAACCCCTACACCATGACCTCCGCTTCATCGCCGCGGACACCTTCGAGAAGCTC 378  
QY 81 AlaValLeuAlaLeuLeuAlaLeuAlaSerArgGlyLeuSerSerProArgAlaLeuGly 100  
DB 379 ATCGTCTCGCGTCTCTCACCTGTGGAGCCAC---CTCTCC-----CGCGCGGCTCC 429  
QY 101 LeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThrLeuValMetGly 120  
DB 430 CTGAGTGGACCATCACCTCTTCTCCCTTCACGCTGCCCAACACGCTCGTCATGGGG 489  
QY 121 IleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeuMetValGlnVal 140  
DB 490 ATCCCGCTGCTGAAGGGGATGTACGGGGAGTTCTCC---GGTAGCCTCATGTGTGCAGATC 546  
QY 141 ValValLeuGlnCysIleIleIleTptTyrThrLeuMetLeuPheLeuPheGluTyrArgAla 160  
DB 547 GTGGTGTCTCAGTGCATCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 606  
QY 161 AlaArgAlaLeuValLeuAspGlnPheProAsp-----GlyAlaAlaAlaSerIleVal 178  
DB 607 GCCAGGATCTCATCCCGGAGGATTTCCCGGACACCGCGGGCGCCATCGCTCCATCGTC 666  
QY 179 SerPheArgValAspSerAspValValSerLeu-----AlaArgGlyAspValGluLeu 196  
DB 667 -----GTGACCCCGACGTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 717  
QY 197 GluAlaGluProAspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGlyAspAla 216  
DB

```
Db 718 GAGGCCGAGGTGAG-----CAGGAC 738
Qy 217 GlyArgValArgValThrValArgLysSerThrSerSerArgSerGluAlaAalaCysSer 236
Db 739 GGCAGATACACGTCACTCCGCGCCCTCCAAAGCGGTGCGCTCGGAGCTTCTACTCGCGG 798
Qy 237 HisSerHisSer--GlnThrMetGlnProArgValSerAsnLeuSerGlyValGluIle 255
Db 799 CGTCTCATGGGTTTCTCCAGCACCGCCGCGCGGAGCACTCACCAACGCCGAGATC 858
Qy 256 TyrSerLeuGlnSerSerArgHsnProThrProArgGlySerSerPheAsnHisAlaAsp 275
Db 859 TACTCGCTGCGTGTGCGCAACCGACGCGCGCGGCTCCAGCTTCAACACACTGAC 918
Qy 276 PhePheAsnLleValGlyAlaAlaLysGlyGlyGlyValAlaAlaGlyAsp----- 293
Db 919 TTCTACTCATGTGTCGG-----CGCAGCTCAACTTCGCCCGCGGAGCCGCTTC 969
Qy 294 -----GluGluLysGlyAlaCysGly 300
Db 970 GGGGTGCGCACCGCGCCACCGCGCGCGCTCCAACTACGAGGAGGACGCGCGCGCGCC 1029
Qy 301 GlyGlyGlyGly-----GlyHisSerProGlnProGln--AlaValAlaValPro 316
Db 1030 AACAGCGCGCGCAGCAAGTACGGGCGAGTACCGCGCGCCCAACCCGCGCATGCGCGCGG 1089
Qy 317 AlaLys-----ArgLysAspLeuHis 323
Db 1090 CCCAACCCCAAGAGCGGCCCAACGCGCAGCCAGCCAGGCGGAGGAGCGCAAGACCTCCAC 1149
Qy 324 MetLeuValTrpSerSerAlaSerProValSerGluArgAlaAlaValHisValPhe 343
Db 1150 ATGTTCTGTGGAGCTCCAGCGCGTGTGCGCGTGTGCGAC-----GTCTTC 1194
Qy 344 GlyValaGlyValaAlaAspHisAlaAspValLeuAla----- 355
Db 1195 GGC---AACGGCGCGAGTACACGACGCGCGCGCTCAAGAGGTCGCGCATGCGCGTC 1251
Qy 356 LysGlyAlaAlaAlaTyrAspGluTyrGlyArgAspTyrSerSer----- 371
Db 1252 GCCTCGCGCGCAAGCGGACCGCGTGGAGAGGAGGACACTTCAGCTT-CGGGAACAGAGG 1310
Qy 372 ArgThrLysAsnGlySerGlyGly-AlaAspLysGlyGlyProThrLeuSerLysLeuG1 391
Db 1311 CGTGGCGGAGGAGCGGAGCGCGCGGACGAGAGAGCGGTGGCGCGCGCGGTGCGG 1370
Qy 391 YSerAsnSerThrAlaGlnLeuTyrProLysAspAspGlyGlyArgAlaAlaAlaVal 411
Db 1371 TGAGCATGGCAAGCCTGGCTTGACGCG-----GCGCGGAC 1406
Qy 411 AlaMetProProAlaSerValMetThrArgLeuLleLeuLleMetValTrpArgLysLe 431
Db 1407 GCGGATGCGCGCCAGCAGCGGTGATGACGCGGCTCATCTCGTATCATGTTGGCGCAAGCT 1466
Qy 431 uileArgAsnProAsnThrTyrSerSerLeuLleGlyValValTrpSerLeuValSerTy 451
Db 1467 CATCGCAACCCCAACACTCTCCAGCTCATCGGCTCATCTGCTGCTGCTGCT 1526
Qy 451 rArgTrpGlyLleGluMetProAlaLleAlaArgSerLleSerLleSerAspAl 471
Db 1527 CCGTGGAACTTCGAGATGCGCGCATCATCTCTGAAATCCATCTCCATCTCTCGACGC 1586
Qy 471 aGlyLeuGlyMetAlaMetPheSerLeuGlyLeuPheMetAlaLeuGlnProArgLle 491
Db 1587 CGGCTCGGATGCGCATGTTTCAGCTTCGCGCTGTTTCATGGCGTGCAGCGCGGATCAT 1646
Qy 491 eAlaCysGlyAsnLysLeuAlaAlaLleAlaMetGlyValArgPheValAlaGlyProAl 511
Db 1647 CGCGTGGGGAACAAGGTGGCGAGCTTCGCGCATGCGGCTGCGGTTCCTGACGCGCGCGC 1706
Qy 511 aValMetAlaAlaSerLleAlaValGlyLeuArgGlyValLeuLeuHisLleAla 531
Db 1707 CGTCATGGCGCGCGCTTCATGCGCGTTCGCGCTTCGCGGCGACGCTCTCTCAAGTGCCTAT 1766
```

```
Qy 531 eValGlnAlaAlaLeuProGlnGlyLleValProPheValPheAlaLysGluTyrGlyVal 551
Db 1767 CGTGCAGCAGCGTGTGCCACAGGCGATGTGCCCCCTTCGCTTCGCCAGAGGTACAGCGT 1826
Qy 551 LHisProAspLleLeuSerThrAla-----TyrGly-----ProIleth 564
Db 1827 GCACCCCGACATCTCAGCAGCGCGGTCACTTTGGCATGCTCATCGCGTGCCTCATC 1886
Qy 564 r 564
Db 1887 G 1887

RESULT 14
AX654121 1827 bp DNA linear PAT 22-MAR-2003
LOCUS AX654121
DEFINITION Sequence 3991 from Patent WO03000898.
ACCESSION AX654121
VERSION AX654121.1 GI:29156935
KEYWORDS
SOURCE
ORGANISM
Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzae; Oryza.
REFERENCE
1 Chang,H.S., Chen,W., Cooper,B., Glazebrook,J., Goff,S.A., Hou,Y.M.,
Katagiri,F., Quan,S., Tao,Y., Whitham,S., Xie,Z., Zhu,T. and Zou,G.
Plant genes involved in defense against pathogens
Patent: WO 03000898-A 3991 03-JAN-2003;
Syngenta Participations AG (CH)
FEATURES
source
1. 1827
/organism="Oryza sativa"
/mol_type="unassigned DNA"
/db_xref="taxon:4530"
ORIGIN

Alignment Scores:
Pred. No.: 3,86e-84 Length: 1827
Score: 1647.50 Matches: 374
Percent Similarity: 66.09% Conservative: 45
Best Local Similarity: 58.99% Mismatches: 100
Query Match: 56.62% Indels: 115
DB: 6 Gaps: 18

US-10-030-884-14 (1-573) x AX654121 (1-1827)
Qy 11 LeuThrAlaValValProLeuTyrValAlaMetThrLeuAlaTyrGlySerValArgTrp 30
Db 1 ATGGCGCGGTGTGTGCGGTGTACGTGGCGATGTTCCTGGCGTACGCGTGGTGGTGG 60
Qy 31 TrpArgLlePheThrProAspGlnCysSerGlyLleAsnArgPheValAlaLeuPheAla 50
Db 61 TGGGGCATCTTCACGCGGACCAAGTGTCTCGGCATCAACCGCTTCGTGCGCATCTTGGCC 120
Qy 51 ValProLeuLeuSerPheHisPheIleSerThrAsnAspPropheAlaMetAsnLeuArg 70
Db 121 GTGCGGCTCTGTCTCTCCATCTCTCCACCAACGACCGGTACGCGCATGAACTCCGC 180
Qy 71 PheLeuAlaAlaAspThrLeuGlnLysValAlaValLeuAlaLeuAlaLeuAlaSer 90
Db 181 TTCCTGGCGGCGGACACGCTGCAGAGAGTGTCTGTCTGCGGGGCTCGCGCGGTGTCG 240
Qy 91 ArgGlyLeuSerSerProArgAlaLeuGlyLeuAspTrpSerLleThrLeuPheSerLeu 110
Db 241 CGCTCCCTCGCGGACCGGCGCGCGGTGGACTGGTCCATCAGCTCTTCTCTCCCTC 300
Qy 111 SerThrLeuProAsnThrLeuValMetGlyLleProLeuLeuArgGlyMetTyrGlyAla 130
Db 301 TCCAGCTGCCCAACACGCTGTATGGGATCCCGTGTGTATCGCATGTGAGCGGCA 360
Qy 131 SerSerAlaGlyThrLeuMetValGlnValValLeuGlnCysLleIleIleTrpTyrThr 150
```



[illegible]

Db 1336 GTATGACAGCCTCATCTCCATAAATGGTGTGGCGCAAGTCATCCGCACACCACCAACT 1339

Qy 438 TyrSerSerLeuIleGlyValValTrpSerLeuValSerTyrArgTrpGlyIleGlutMet 457  
:  
Db 1396 TATTCCAGCCTCTCGGCCCTCGCGCTGGTGCTCTCTCGCTTTCCGCTGTCACGTTCCATC 1455

Qy 458 ProAlaIleAlaLargSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMet 477  
:  
Db 1456 CAGCAATGTCGAGAAGTCAATCTCCATTTCTTCGNACGACAGCCTGGGATGGCCATG 1515

Qy 478 PhSeuLeugLyLeuPheMetAlaLeuGlnProArgIlellelaCySGLysLeu 497  
:  
Db 1516 TTTAGCCTGGGAATTGTTTCATGGCGCTCGACCCAGCATCATCGCTGTGCACAATCAGCC 1575

Qy 498 AlaAlaIleAlaMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaLasr 517  
:  
Db 1576 GCGTGCTCTCCAATGGCGGCTTCCTCGGGGCCCTCGCTCATCGCGCGCGCTCA 1635

Qy 518 IleAlaValGlyLeuArgGlyValLeuLeuHisIleAlaIleValGlnAlaLeuPro 537  
:  
Db 1636 ATGCCATCGGACTCCGCGGAGCGCTCTCGACGTGCCATTGTTTCAGGGGGCTCTACCA 1695

Qy 538 GlnGlyIleValPropheValPheAlaIleGluTyrclyValHisProAspileLeuSer 557  
:  
Db 1696 CAAGGGAAATGTGCTTTGTGTTTTGCAAAAAGATAACAATGTCCACCGGCCATCTCGAGC 1755

Qy 558 ThrAla-----TyrGly-----Proilethr 564  
:  
Db 1756 ACAGCGTAATTTTGGCATGCTAATAGCTTCTCCAATCAC 1797

RESULT 15  
AY093960 1869 bp mRNA linear PLN 27-APR-2002  
LOCUS Arabidopsis thaliana Atlg73590/F6D5\_2 mrna, complete cds.  
DEFINITION Arabidopsis thaliana  
ACCESSION AY093960  
VERSION AY093960.1 GI:20334719  
KEYWORDS FLU CDNA.

SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi s.  
1 (bases 1 to 1869)

REFERENCE Chouk,R., Chen,H., Kim,C.J., Meyers,M.C., Shinn,P., Banh,J., Bowser,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.  
Arabidopsis ORF clones

TITLE Unpublished  
JOURNAL (bases 1 to 1869)  
REFERENCE Chouk,R., Chen,H., Kim,C.J., Meyers,M.C., Shinn,P., Banh,J., Bowser,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.  
Arabi dopsis ORF clones

TITLE Direct Submission  
JOURNAL Submitted (02-APR-2002) Salk Institute Genomic Analysis Laboratory (SIGnAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

COMMENT RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL CDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai.T., Carninci.P., Kawai,J., Hayashizaki.Y. and Shinozaki.K.



The Salk, Stanford, FPEC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAFL cDNAs: Cheuk, R., Chen, H., Kim, C. J., Meyers, M. C., Shinn, P., Banh, J., Bowser, L., Chan, M. M., Chang, E., Dale, J. M., Deng, J. M., Goldsmith, A. D., Jones, T., Karlin-Neuman, G., Lam, B., Lee, J. M., Lin, J., Miranda, M., Nguyen, M., Onodera, C. S., Palm, C. J., Quach, H. L., Southwick, A., Tang, C. C., Toriumi, M., Wu, H. C., Yamada, K., Yamamura, Y., Yu, G., Davis, R. W., Theologis, A., and Ecker, J. R.

Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J. R. (SSP/Salk) contributed equally to this work as PIs.

FEATURES  
source Location/Qualifiers  
1. .1869  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="1"  
/clone="U12338"  
/ecotype="Columbia"  
/notes="This clone is in pUNI 51"  
CDS  
1. .1869  
/notes="auxin transporter splice variant b, putative"  
/codon\_start=1  
/product="Atlg73590/F6D5\_2"  
/protein\_id="AAM16221.1"  
/db\_xref="GI:20334720"  
/translation="MITRADFVHTVMTVPLVYVIMLAVGVKWKIPTPQCQGINR  
FVALPVLPLLSHFIAANPYANFLAASLQKVIIVLSLFLWCKLSRNGSLDWI  
TFLSLTPTNLTIVGLIKMGYGNFSDLMYQIVVLCIIVYTLMLFLFYRGAKLL  
ISQPPDTAGSIVSIVHSDIMSLDGRQPLEATEIKEDKLHVTVRSNASRSDIYS  
RRSQGLSATPRSNLTNABIIYSLOSRRNTPRSGSFNHTDFYSWASGGGRNSNFGPG  
EAVFGSGKTPRPSNVEEYDGGPAKTAAGTAAGREHYQSGSGGGGGAHPVAPNG  
MPSNPTGGGGTAAGNA PVVGGKQDNGRDLHFMVSSSASPVSVDVFGGGGNHHA  
DYSTATNDHQVKIISVPOGNSNDQYVEREESFGNKKDDSKVLATDGGNNISKTT  
QAKVMPPTSVTLILIMLWKLIRNPNSISLFGITWLSIFKNWIEPALIAKSIIS  
ILSDAGLGMFSLGIFMALNPRIIACGNRRFAAFAMRFVVGPAVMLVASYAVGLRG  
VLLHVAIIQAALPQGIVPFVFAKEYNVHPDILSTAVIFGMLIALPILLYILLGI"

ORIGIN

Alignment Scores:  
Pred. No.: 3,96e-84 Length: 1869  
Score: 1647.50 Matches: 366  
Percent Similarity: 64.82% Conservative: 56  
Best Local Similarity: 56.22% Mismatches: 104  
Query Match: 56.62% Indels: 125  
DB: 8 Gaps: 18  
US-10-030-884-14 (1-573) x AY093960 (1-1869)

Qy 1 MetileThraAlaLeuAspLeuTyHisValLeuThraAlaValValProLeuTyValAla 20  
Db 1 ATGATTACGGCGGGGACTTCTACCGCTTATGACGGCTATGGTCCGTATACGTAGCT 60  
Qy 21 MetThrLeuAlaTyGlySerValArgTTPArgIlePheThrProAspGlnCysser 40  
Db 61 ATGATCTCTCGCTACGGCTCTGCAATGTGTGGAATCTTACACACACCAATGCTCC 120  
Qy 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
Db 121 GGCTAATACCGTTTGGTCTCTTCTCGCGCTTCTCTCTCTCTCTCTCTCTCTCTCT 180  
Qy 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnIlyVal 80  
Db 181 GCTAAACACCGTTACGGCATGACCTCGTTTCTCGCCGCGACGATTCTCTCCAGAAAGTC 240  
Qy 81 AlaValLeuAlaLeuAlaLeuAlaLeuAlaLeuAlaLeuAlaLeuAlaLeuGly 100  
Db 241 ATTGCT 291  
Qy 101 LeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThrLeuValMetGly 120  
Db 292 TTAGATTGGACCAAACT 351

Qy 121 IleProLeuLeuArgGlyMetTyGlyAlaSerSerAlaGlyThrLeuMetValGlnVal 140  
Db 352 ATACCTCTTCTCAAGGCAATGATGGT---AATTCTTCGGCGACCTCATGTGTTCAATC 408  
Qy 141 ValValLeuGlnCysserIleLeuTTPThrLeuMetLeuPheLeuPheGlyTyArgAla 160  
Db 409 GTTGTCTTTCAGTGATCAITTTGTACACTCATGCTCTTCTCTTGTGATGACCGTGA 468  
Qy 161 AlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSerIleValSerPhe 180  
Db 469 GCTAAGCTTTTGTATCTCCGAGCAGTTTCCAGAC---ACACGAGGATCTATTGTTGAT 525  
Qy 181 ArgValAspSerAspValValSerLeu---AlaArgGlyAspValGluLeuGluAlaGlu 199  
Db 526 CATGTTGATTCGACATTTATGCTTTAGATGGAGACAACTTTTGGAACTGAGCTGAG 585  
Qy 200 ProAspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGlyAlaGlyArgVal 219  
Db 586 ATTTAA-----GAAGATGGAGAGCTT 606  
Qy 220 ArgValThrValArgLysSerThrSerArgSerGluAlaAlaCysSerHisSerHis 239  
Db 607 CATGTTACTGTCGTCGTTCTAATGCTTCAAGGCTGATATTTACTCGAAGAGTCTCAA 666  
Qy 240 SerGlnThrMetGlnProArgValSerAsnLeuSerGlyValGluIleTySerLeuGln 259  
Db 667 GGCTTATCTCGACACCTAGACCTTCGAATCTAACCAACGCTGAGATATATTCGCTTCA 726  
Qy 260 SerSerArgAsnProThrProArgGlySerSerPheAsnHisAlaAspPheAsnIle 279  
Db 727 AGTTCAAGAAACCCACGCCACGCTGCTAGTTTTAACTATCATGATGATTTTACTCGATG 786  
Qy 280 Val----- 280  
Db 787 ATGCTCTCTGCTGGTGGTGGAACTCTAACTTTGGTCTCGAAGAGCTGTGTTGTTGTTCT 846  
Qy 281 -----GlyAlaAlaAlaLys----- 285  
Db 847 AAGGTCCTACTCTCGAGACCTTCCAACTACGAAAGACGGTGGTCTGCTGTAACCGCAG 906  
Qy 286 GlyGlyGlyGlyAlaAlaGly-----AspGluGluLysGlyAlaCysGlyGly 301  
Db 907 CTGCTGGAAGTCTGCTGGAGCTGGGAGGTTTCAATATCAATCTGGAAGAGTGGTGGC 966  
Qy 302 GlyGlyGlyGlyHisSerProGlnPro----- 310  
Db 967 GTTGAGGAGCGCATTTATCCGCGCGCCGAAACCCAGGGATGTTTTCGCCCAACACTGGCG 1026  
Qy 311 -----GlnAlaValAlaValProAlaLysArg----- 319  
Db 1027 GTTGAGGAGCGCGCGCGGAAAGGAAACGCTCCGGTGGTGGTGGGAAAGACAGACGGA 1086  
Qy 320 -----LysAspLeuHisMetLeuValTrpSerSerSerAlaSerProValSerGluArg 337  
Db 1087 AACGGAAGAGATCTTTCATGTTTGTGGAGCTCAAGTGTCTTCGCGGTCTTCAGAT--- 1143  
Qy 338 AlaAlaValHisValPheGlyAlaGlyAlaAlaAspHisAlaAspValLeuAlaLysGly 357  
Db 1144 -----GTGTTCCGCGGTGGAGAGGAAACCCACAC-----GCCGATTAC 1182  
Qy 358 AlaGlnAlaTyAspGluTyGlyArgAsp----- 367  
Db 1183 TCCACCGCTACGAACGATCATCAAAAGACGTTAAGATCTCTGTACCTCAGGGGAATAGT 1242  
Qy 368 -----AspTySerSerArgThrLysAsnGlySerGlyGlyAlaAspLysGlyGly 384  
Db 1243 AACGACAAACAGTACGTGGAGAGGAGGTTTGTGTTTCGTTAAACAAGACGATGATAGC 1302  
Qy 385 ProThrLeuSerLysLeuGly-----SerAsnSerThrAlaGlnLeuTyPro 400  
Db 1303 AAAGTATTGGCAACCGGCGGTGGGAAACACATAGCAACAAACACGACGAG----- 1353

```
Qy 401 LysAspAspGlyGluGlyArgAlaAlaAlaValAlaMetProProAlaSerValMetThr 420
Db 1354 -----GCTAAGGTGATGCCACCACCAAGTGTGATGACA 1386
Qy 421 ArgLeuIleLeuIleMetValTIPArgIleLeuIleArgAsnProAsnThrTyrSerSer 440
Db 1387 AGACTCATTCTCATATGGTTTGGAGAACTTATTGTAATCCCACTCTTACTCCAGT 1446
Qy 441 LeuIleGlyValValTIPSerLeuValSerTyrArgTIPGlyIleGluMetProAlaIle 460
Db 1447 TTATTGGCATCACCTGCTCCTCATTTCTTCAAGTGAACATTGAAATGCCAGTCTTT 1506
Qy 461 IleAlaArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeu 480
Db 1507 ATAGCAAGTCTATCTCCATCTCTCAGATCCAGGTTTAGCATGGCTATGTTCACTCTT 1566
Qy 481 GlyLeuPheMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAlaAlaIle 500
Db 1567 GGGTTGTTTCATGGCGTTAAACCCCAAGAAATAATAGCTTGTGAAACAGAGAGAGCTTTT 1626
Qy 501 AlaMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaAlaSerIleAlaVal 520
Db 1627 GCGCGGCTATGAGATTGTGTTGGACCTGCCCTCATGCTCATGCTGTTCTTTATGCCGTT 1686
Qy 521 GlyLeuArgGlyValLeuLeuHisIleAlaIleValGlnAlaAlaLeuProGlnGlyIle 540
Db 1687 GGCCTCCGTCGCGTCTCTCCATGTTGCCATTATCCAGGCAGCTTTGCCCAAGGAATA 1746
Qy 541 ValProPheValPheAlaLysGluTyrGlyValHisProAspIleLeuSerThrAla--- 559
Db 1747 GTACCGTTTGTGTTTGCCAAAGAGTATAATGTGCATCTCTGACATTCTTAGCACTGCGGTG 1806
Qy 560 ---TyrGly-----ProIleThr 564
Db 1807 ATATTGGGATGTTGATCGCGTTCGCCATAACT 1839
```

Search completed: March 3, 2004, 11:11:39  
Job time : 6430 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 3, 2004, 09:12:55 ; Search time 3990 Seconds  
(without alignments)  
4288.479 Million cell updates/sec

Title: US-10-030-884-14

Perfect score: 2910

Sequence: 1 MITALDLIYHLTVAVPLYVA.....DILSTAYGPTSGFITCS 573

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xih  
-Q=/cgn2 1/USPT0.spool/US10030884/runat 24022004 103923 1960/app query.fasta\_1.711  
-DB=EST -QFMT=fastcap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DICALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10030884 @cgn 1 1906 @runat 24022004 103923 1960 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG SCORE=0 -WAIT -DSFBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hcti:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hcti:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_eston:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mus:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1922	66.0	2737	11 AY110494	AY110494 Zea mays
2	1405	48.3	840	29 CG358070	CG358070 OGBK16TV
3	1356	46.6	893	29 CG376782	CG376782 OGBK16TV
4	1222	42.0	835	29 CC722061	CC722061 OGUZ02TV
5	1204	41.4	730	29 CG376770	CG376770 OGUZ02TV
6	1100	37.8	776	29 CG317974	CG317974 OGWAG1TV
7	1043.5	35.9	658	13 BU098540	BU098540 946136B05
8	1008.5	34.7	627	13 BU037419	BU037419 946138F12
9	1004	34.5	426	11 AY106004	AY106004 Zea mays
10	999.5	34.3	610	13 BU080190	BU080190 946150G03
11	997.5	34.3	604	14 CA831110	CA831110 1117015F0
12	983	33.8	630	14 CA830783	CA830783 1117011D1
13	962	33.1	726	29 CG358058	CG358058 OGBK16TH
14	936	32.2	1221	11 AY110024	AY110024 Zea mays
15	909	31.2	972	28 CG371166	CG371166 PURSK16TB
16	907	30.4	903	14 CF244269	CF244269 3530_1_28
17	885.5	30.4	903	14 CF243672	CF243672 3530_1_23
18	881.5	30.3	552	13 BU037594	BU037594 946140E03
19	848.5	29.2	1105	14 CK208792	CK208792 FGAS02051
20	837	28.8	1107	14 CK209475	CK209475 FGAS02124
21	836.5	28.7	1109	14 CK208849	CK208849 FGAS02057
22	819.5	28.2	739	12 BG600488	BG600488 EST505383
23	809.5	27.8	868	29 CG317965	CG317965 OGWAG11TH
24	805.5	27.7	497	13 BU036965	BU036965 946131B07
25	800.5	27.5	760	29 CC722052	CC722052 OGUZ02TV
26	798.5	27.4	728	14 CB348868	CB348868 CAB2SG000
27	791.5	27.2	796	9 AF162834	AF162834 AF162834
28	778.5	26.8	774	14 CF437477	CF437477 EST7673822
29	769	26.4	963	29 CG861094	CG861094 ZMMB0026
30	762	26.2	619	14 CA236508	CA236508 SCSEBFL502
31	758	26.0	819	14 CB648277	CB648277 OSUNEB11J
32	752	25.8	741	13 BQ999371	BQ999371 OGUZ1008
33	748	25.7	861	14 CK279869	CK279869 EST725947
34	748	25.7	897	14 CK282170	CK282170 EST744892
35	745.5	25.6	931	14 CK293778	CK293778 EST756492
36	733.5	25.2	851	28 BH553421	BH553421 BOGDR35TF
37	731.5	25.1	702	28 BZ011975	BZ011975 Oe386B11
38	730	25.1	469	14 CD001628	CD001628 3529_1_10
39	729	25.1	696	14 CA184766	CA184766 SCSEBFL309
40	727.5	25.0	633	14 CF805982	CF805982 psHB002x1
41	727.5	25.0	929	14 CK294982	CK294982 EST757696
42	724.5	24.9	739	14 CB646582	CB646582 OSUNEB09C
43	720.5	24.8	913	14 CK250728	CK250728 EST734365
44	718.5	24.7	901	14 CK292860	CK292860 EST755574
45	717.5	24.7	894	14 CK284513	CK284513 EST747235

# ALIGNMENTS

RESULT 1  
AY110494  
LOCUS AY110494 Zea mays CL464\_-1 mRNA 2737 bp linear HTC 17-OCT-2002  
DEFINITION Zea mays CL464\_-1 mRNA sequence.  
ACCESSION AY110494  
VERSION AY110494.1 GI:21214903  
KEYWORDS HTC.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 2737)

AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whittitt,M.S.,  
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.  
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of  
Overgo Probes  
JOURNAL Unpublished (2002)  
AUTHORS 2 (bases 1 to 2737)  
TITLE Direct Submission  
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of  
Missouri, Columbia, MO 65211, USA  
COMMENT These are publicly available from ZmDB and may be found by BLAST  
searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,  
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the  
maize cDNA sequences is either Virginia Walbot, Stanford or Pat  
Schnable, Iowa State, then clones may be requested from ZmDB:  
www.zmdb.iastate.edu.

## FEATURES

source

Location/Qualifiers  
1..2737  
/organism="Zea mays"  
/mol\_type="mRNA"  
/db\_xref="MaizeDB:632106"  
/db\_xref="taxon:4577"  
/clone\_lib="Maize Mapping Project/DuPont Consensus  
Library"  
/note="this sequence is part of a project of EST  
assemblies resulting from the application of public  
contigs to seed Dupont contigs; this resource was  
assembled by Dupont as part of a collaboration for the  
overgo addressing of BACs in conjunction with the Maize  
Mapping Project"

## ORIGIN

## Alignment Scores:

Pred. No.: 4,75e-135 Length: 2737  
Score: 1922.00 Matches: 406  
Percent Similarity: 70.98% Conservative: 0  
Best Local Similarity: 70.98% Mismatches: 166  
Query Match: 66.05% Indels: 0  
DB: 11 Gaps: 0

US-10-030-884-14 (1-573) x AY110494 (1-2737)

QY 1 MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValProLeuTyrValAla 20  
DB 171 ATGATCACCGCGTGGACCTCTACACAGTGTGACGGCGGTGGTCCGCTGTACGTGCC 230  
QY 21 MetThrLeuAlaTyrGlySerValArgTTPArgIlePheThrProAspGlnCysSer 40  
DB 231 ATGACGCTGGCGTACGGCTCCGTCCGCTGGTGGCGCATCTTCACGCGGACGAGTGCTCC 290  
QY 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
DB 291 GGGATCAACCGCTTCTGGCGGCTCTTCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 350  
QY 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAspThrLeuGlnLysVal 80  
DB 351 NCCAACGACCCCTTCCCATGACCTGCGCTCTCTGGCCGCGCACGCTGCAGAGGTG 410  
QY 81 AlaValLeuAlaLeuLeuAlaLeuAlaSerArgGlyLeuSerSerProArgAlaLeuGly 100  
DB 411 GCCGCTCTCGCGTCTGCGCGCTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 470  
QY 101 LeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThrLeuValMetGly 120  
DB 471 NNNNNCTGGAGCATCATCGCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCA 530  
QY 121 IleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeuMetValGlnVal 140  
DB 531 ATCCCGCTGTGGAGGATGATACGGCGGTGCTCGCGCGGACCGCTCATGTGTCCAGNN 590  
QY 141 ValValLeuGlnCysIleIleTrpTyrThrLeuMetLeuPheLeuPheGluTyrArgAla 160

DB 591 NNN 650  
QY 161 AlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAspSerIleValSerPhe 180  
DB 651 NNN 710  
QY 181 ArgValAspSerAspValValSerLeuAlaArgGlyAspValGluLeuGluAlaGluPro 200  
DB 711 NNN 770  
QY 201 AspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGlyAspAlaGlyArgValArg 220  
DB 771 GNN 830  
QY 221 ValThrValArgLysSerThrSerSerArgSerGluAlaAlaCysSerHisSerHisSer 240  
DB 831 NNCACGCTGGCGAAGTCCACAGCTCGCGCTCGAGCGCGCGTCTCGCAGCTCGCAGCTCC 890  
QY 241 GlnThrMetGlnProArgValSerAsnLeuSerGlyValGluIleTyrSerLeuGlnSer 260  
DB 891 CAGACCATGACAGCCCGCTGTGTCCAACTCTCGGGCTGGAGATCTACTCGCTGCAGTCG 950  
QY 261 SerArgAsnProThrProArgGlySerSerPheAsnHisAlaAspPhePheAsnIleVal 280  
DB 951 NNN 1010  
QY 281 GlyAlaAlaAlaLysGlyGlyGlyAlaAlaGlyAspGluGluLysGlyAlaCysGly 300  
DB 1011 GGGCGCCCGCNGNN 1070  
QY 301 GlyGlyGlyGlyHisSerProGlnProGlnAlaValAlaValProAlaLysArgLys 320  
DB 1071 NNN 1130  
QY 321 AspLeuHisMetLeuValTTPSerSerSerAlaSerProValSerGluArgAlaAlaVal 340  
DB 1131 GACCTGCACATGCTCTGTGGAGCTCCAGCGCTCCCGCTGTCGAGCGCGCGCGCGTGG 1190  
QY 341 HisValPheGlyAlaGlyGlyAlaAspHisAlaAspValLeuAlaLysGlyAlaGlnAla 360  
DB 1191 CACGCTCTTCGGCGCGCGCGCTGACCATGCGCGCTCTCGCCAAAGAGAGCCAGGCC 1250  
QY 361 TyrAspGlnTyrGlyArgAspAspTyrSerSerArgThrIleAsnGlySerGlyAla 380  
DB 1251 TACGACGAGTACGGCGCGCGAGCTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1310  
QY 381 AsplysGlyGlyProThrLeuSerLysLeuGlySerAsnSerThrAlaGlnLeuTyrPro 400  
DB 1311 NNN 1370  
QY 401 LysAspAspGlyGluGlyArgAlaAlaAlaValMetProAlaSerValMetThr 420  
DB 1371 NNN 1430  
QY 421 ArgLeuIleLeuIleMetValTTPArgLysLeuIleArgAsnProAsnThrTyrSerSer 440  
DB 1431 CGGCTCATCTCATCTCATCTGTGTGGAGAGAGTGTATCCGGAACCCCAACACCTACTCCAGC 1490  
QY 441 LeuIleGlyValValTTPSerLeuValSerTyrArgTTPGlyIleGluMetProAlaIle 460  
DB 1491 CTATCGCGCTCGCTCTGCTGCTCCCTGCTCTCTA CAGGTGGGCGCATCGAGATCCAGGATC 1550  
QY 461 IleAlaArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeu 480  
DB 1551 ATCGCCCGGTCGATTTGATCTCTGTGGACCGCGGTCTCGGGATGGCCATGTTACGCTTA 1610  
QY 481 GlyLeuPheMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAlaIle 500  
DB 1611 GGCCTGTTTCATGCGCTGCAGCCGAGGATCATCGCTGGGGAAACAAAGCTGGCGGCGCATC 1670  
QY 501 AlaMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaAlaSerIleAlaVal 520  
DB 1671 GCGATGGCGCTCGGTTCTGTCGAGGCGCGCGGTTCATGGCCCGCGCTCCATCGCGCTC 1730

```

Qy 521 GlyLeuArgGlyValLeuLeuHisLeuValLeuAlaLeuProGlnGlyLeu 540
Db 1731 GGTCTGGCGGCGTCTCTCCATCGCCATCGTCCAGGCTGCTGCTCAGGGGATC 1790
Qy 541 ValProPheValPheAlaLysGlyTyrGlyValHisProAspIleLeuSerThrAlaTyr 560
Db 1791 GTCCCGTTCGTGTCGCAAGAGTACGGGCTTCATCCCGACATCTCGACACAGCGTAT 1850
Qy 561 GlyProIleThrSerHisGlyPheIleThrCysHis 572
Db 1851 GGTCCATATACATCGCATGTTTCATCATTCGTCAT 1886

RESULT 2
CG358070 940 bp DNA linear GSS 26-AUG-2003
LOCUS OGI8K16TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBma0724C07,
DEFINITION genomic survey sequence.
ACCESSION CG358070
VERSION
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 840)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other_GSSs: OGI8K16TH
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES
source
1..840
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMBma0724C07"
/clone_lib_ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Alignment Scores:
Pred. No.: 1,53e-96 Length: 840
Score: 1405.00 Matches: 279
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 48.28% Indels: 0
DBs: 29 Gaps: 0

US-10-030-884-14 (1-573) x CG358070 (1-840)

Qy 74 AlaAspThrLeuGlnLysValAlaValLeuAlaLeuAlaLeuAlaSerArgGlyLeu 93
Db 838 GCCGACACGCTGCAGAGGTGCCGCTCTCGCGTCTGCGGCTGCGCTCCCGGCGCTC 779
Qy 94 SerSerProArgAlaLeuGlyLeuAspTrpSerIleThrLeuPheSerLeuSerThrLeu 113
Db 778 TCCCTCCCGCGCGCTCGGGCTCGACTCGAGCATCAGCTCTTCTCCCTCCACGCTC 719
Qy 114 ProAsnThrLeuValMetGlyLeuProLeuLeuArgGlyMetTyrGlyAlaSerSerAla 133

```

```

Db 718 CCCAACACGCTGTCATGGGCATCCCGCTGCTGGAGGCATGTACGGCGCGCTGTCGGCC 659
Qy 134 GlyThrLeuMetValGlnValValLeuGlnCysIleIleIleTyrThrLeuMetLeu 153
Db 658 GGCACGCTCATGGTCCAGTCTGCTGCTCCAGTGCATCACTGGTACACGCTCATGCTC 599
Qy 154 PheLeuPheGluTyrArgAlaAlaArgAlaLeuValLeuAspGlnPheProAspGlyAla 173
Db 598 TTCTCTTCGAGTACCGCGCGCGCGCTGCTGCTCGACCACTTCCCGACGCGGCC 539
Qy 174 AlaAlaSerIleValSerPheArgValAspSerAspValValSerLeuAlaArgGlyAsp 193
Db 538 GCGCGTCCATCTCTCTTCCGCTGACTCCGACGCTGCTGCTCGCCAGGGGGAC 479
Qy 194 ValGluLeuGluAlaGluProAspClyValAlaGlyAlaGlyAlaValSerSerArgGly 213
Db 478 GTGAGCTCGAGCGCGACCGCGCTGCGCGCGCGCGCGCTCTCTCCCGCGGC 419
Qy 214 GlyAspAlaGlyArgValArgValThrValArgLysSerThrSerSerArgSerGluAla 233
Db 418 GGGGACGCGCGCGGTGCGGTCACTCCGCAAGTCCACAGCTCGCTCGCTCCGAGGCC 359
Qy 234 AlaCysSerHisSerHisSerGlnThrMetGlnProArgValSerAsnLeuSerGlyVal 253
Db 358 GCGTCTCGCACTCCGCACTCCAGACCATGAGCCCGCTGTGTCCAACTCTCCCGCGCTG 299
Qy 254 GluIleTyrSerLeuGlnSerSerArgAsnProThrProArgGlySerSerPheAsnHis 273
Db 298 GAGATCTACTCGTCTCAGTCTGTGGCGCAACCCCGCGCGGTTCAGCTTCAACAC 239
Qy 274 AlaAspPheAsnIleValGlyAlaLysGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 293
Db 238 GCCGACTTCTTCAACATGCTCGCGCGCGCGCGCAAGGGAGCGGAGGCGCGGGGAC 179
Qy 294 GluGluLysGlyAlaCysGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 313
Db 178 GAGGAGAGGGCGCATGTGGCGCGCGCGCGCGGAGAGACACTCGCCGCGCGCGAGCCGCTC 119
Qy 314 AlaValProAlaLysArgLysAspLeuHisMetLeuValTrpSerSerSerAlaSerPro 333
Db 118 GCGGTGCGCGCCACAGAGGAGACCTGCACATGCTGCTGAGAGTCCAGCGCTCGGCC 59
Qy 334 ValSerGluAspAlaAlaValHisValPheGlyAlaGlyGlyAlaAspHisAlaAsp 352
Db 58 GTGTCGAGCGCGCGCTGACGCTTTCGCGCGCGCGCGCGCTGACCATGCCGAC 2

RESULT 3
CG376782 893 bp DNA linear GSS 26-AUG-2003
LOCUS OGI3CN2TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBma0776C20,
DEFINITION genomic survey sequence.
ACCESSION CG376782
VERSION CG376782.1 GI:34294049
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 893)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other_GSSs: OGI3CN2TV
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF

```

```

Class: sheared ends.
FEATURES
    source
        Location/Qualifiers
            1..893
                /organism="Zea mays"
                /mol_type="genomic DNA"
                /strain="B73"
                /db_xref="taxon:4577"
                /clone="ZMMBMA0776C20"
                /clone_lib="ZM 0.7 1.5 KB"
                /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
                methylation filtered genomic DNA library"
ORIGIN
Alignment Scores:
Pred. No.:      8,848-93      Length:      893
Score:          1356.00      Matches:    273
Percent Similarity: 99.64%      Conservative: 1
Best Local Similarity: 99.27%      Mismatches: 0
Query Match:      46.60%      Indels:     1
DB:                29          Gaps:         0
US-10-030-884-14 (1-573) x CG376782 (1-893)
QY 1 MetTlThrAlaLeuAspLeuTyrHisValLeuThrAlaValProLeuTyrValAla 20
DB 69 ATGATCACCGCGCTGGACCTTACACAGTGTGACGGGGTGTGCTCCCTGTAGTGCC 128
QY 21 MetTlThrAlaTyrGlySerValArgTyrTrpArgIlePheThrProAspGlnCysSer 40
DB 129 ATGACGCTGGCTACGGCTCCGTCGCTGGTGGCGCATCTTCACGCGGACGAGTGCTCC 188
QY 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60
DB 189 GGGATCAACCGCTTCGTGGCGCTCTTCGCGCGTGGCGCTCTCTCCCTTCATCTTCCTCC 248
QY 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnIysVal 80
DB 249 ACAACAGACCCCTTGGCCATGACCTTGGCTTCTTGGCGCGGACACGCTGCAGAGGTG 308
QY 81 AlaValLeuAlaLeuAlaLeuAlaSerArgGlyLeuSerSerProArgAlaLeuGly 100
DB 309 GCGGCTCTCGCGCTGTGCGCTGGCTCCCGCGGCTCTCTCCCGCGCGCTCGGG 368
QY 101 LeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThrLeuValMetGly 120
DB 369 CTCGACTGGAGCATACGGCTCTTCTCCCTCTCCAGCTTCCCAACACGCTCGTATGGGC 428
QY 121 IleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeuMetValGlnVal 140
DB 429 ATCCCGCTGCTCGAGGCGATGTACGCGCGTGTGTCGCGCGGACGCTCATGTCCAGGTC 488
QY 141 ValValLeuGlnCysIleIleTyrThrLeuMetLeuPheLeuPheGluTyrArgAla 160
DB 489 GTCGTCTCCAGTGCATCATCTGGTACACGCTATGCTCTTCTCTTCGAGTACCGGCC 548
QY 161 AlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSerIleValSerPhe 180
DB 549 GCGCGCGCGCTGCTCTCGACCATGTTCCCGAGCGCGCGCGCTCCATGCTCTCTTC 608
QY 181 ArgValAspSerAspValValSerLeuAlaArgGlyAspValGluLeuAlaGluPro 200
DB 609 CGCGTGCAGTCCGACGCTGTCTGCTGTCGACAGGGGGAGCTCGAGCTCGAGGCCGAGGCC 668
QY 201 AspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGlyAspAlaGlyArgValArg 220
DB 669 GACGGGGTGGCGCGCGCGCGCTCTCTCCCGCGGGGAGCGCGGGGGGTGGCG 728
QY 221 ValThrValArgIlySerThrSerArgSerGluAlaAlaCysSerHisSerHisSer 240
DB 729 GTACCGTGGCAAGTCCACAGCTCGCGTCCGAGGCGCGGTGCTCGACATCGACATCC 788
QY 241 GlnThrMetGlnProArgValSerAsnLeuSerGlyValGluIleTyrSerLeuGlnSer 260

```

---

```

DB 789 CAGACCATGCGAGCCCGCTGTGTCCAACTCTCGGCGTGGAGATCTACTCGTGCAGTCG 848
QY 261 SerArgAsnProThr-ProArgGlySerSerPheAsnHisAla 274
DB 849 TCGGCAACCCACCCCGCGCGGCTCCAGCTTTCACACGCC 891
RESULT 4
CC722061 835 bp DNA linear GSS 19-JUN-2003
CGUI209TV ZM 0.7 1.5 KB Zea mays genomic clone ZMMBMA0452A18,
genomic survey sequence.
CC722061
CC722061.1 GI:32126837
GSS.
Zea mays
Zea mays
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 835)
White, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Other_GSSs: OGUIZ09TH
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.
Location/Qualifiers
1..835
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0452A18"
/clone_lib="ZM 0.7 1.5 KB"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Alignment Scores:
Pred. No.:      1,218-82      Length:      835
Score:          1222.00      Matches:    245
Percent Similarity: 88.13%      Conservative: 0
Best Local Similarity: 88.13%      Mismatches: 0
Query Match:      41.99%      Indels:     33
DB:                29          Gaps:         1
US-10-030-884-14 (1-573) x CC722061 (1-835)
QY 145 CysIleIleTyrThrLeuMetLeuPheLeuPheGluTyrArgAlaAlaArgAlaLeu 164
DB 2 TGCATCATCTGTGTACACGCTCATGCTCTTCTCTTCAGTACCGCGCGCGCGGCTC 61
QY 165 ValLeuAspGlnPheProAspGlyAlaAlaAlaSerIleValSerPheArgValAspSer 184
DB 62 GTCCTCCACAGTTCCTCCCGAGCGCGCGCGCTCCATCGTCTCTTCCGCGTGCAGTCC 121
QY 185 AspValValSerLeuAlaArgGlyAspValGluLeuGluAlaGluProAspGlyValAla 204
DB 122 GACGTCTCTCGTTCGTCGCGAGGGGGAGCTCGAGCTCGAGCGCGCGCGGCTGCC 181
QY 205 GlyAlaGlyAlaValSerSerArgGlyGlyAspAlaGlyArgValArgValThrValArg 224
DB 182 GCGCGCGCGCGCTCTCTCCCGCGGGGAGCGCGGGGGTGGCGTGCAGTGCAGTGGCG 241
QY 225 LysSerThrSerArgSerGluAlaAlaCysSerHisSerHisSerGlnThrMetGln 244

```

242 AAGTCCACAGCTCGCGCTCGAGGCGCGTCTCGCATCTCGCACTCCAGACCATGAG 301  
245 ProArgValSerAsnLeuSerGlyValGluLeuTyrSerLeuGlnSerSerArgAsnPro 264  
302 CCCCCTGTGTCCACCTCTCGCGCTGGAGATCTACTCTCGCTGCGAGTCGTCGCGCAACCCC 361  
265 ThrProArgGlySerSerPheAsnHisAlaAspPheAsnIleValGlyAlaAla 284  
362 ACCCGCGCGGCTCCAGCTTCAACACGCGGAGCTTCTTCAACATCGTGGCGCGCGCC 421  
285 LysGlyGlyGlyGlyAlaAlaGlyAspGluGluLeuGlyGlyAlaCysGlyGlyGly 304  
422 AAGGAGCGCGAGAGCGCGCGGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 481  
305 GlyHisSerProGlnProGlnAlaValAlaValProAlaValArgLysAspLeuHisMet 324  
482 GGACACTCGCGCGAGCGCGAGCGCGCTCGCGCTCGCGCCCAAGAGGAGGAGCTGCACATG 541  
325 LeuValTrpSerSerAlaSerProValSerGluArgAlaAlaValHisValPheGly 344  
542 CTGCTGGAGCTCCAGCGCTCGCGCTCGCGCTCGAGCGCGCGCTGCACGCTCTTCGCG 601  
345 AlaGlyGlyAlaAspHisAlaAspValLeuAla-Lys----- 356  
602 GCCGCGCGCGCTGACCATGCGAGCTCTCGCGCAAGGTTGTAACATGCGTGATCTTT 661  
356 ----- 356  
662 TGCTGTATTCTACCAAAATGTTGCTGTGCAATTTTATTGTCACCAATGGGTGTG 721  
357 -----GlyAlaGlnAlaTrpAspGluTyrGlyArgAspAspTyrSerSerAr 372  
722 CACAACACAAAGCAGGAGCGCGAGCTTACGAGTACGCGCGCGAGGAGGAGGAGGAGGAG 781  
372 gThrLysAsnGlySerGlyGlyAlaAspLysGlyGlyProThrLeuSerLys 389  
782 GACGAAGAAAGGAGCGCGCGCGCGAGCAAGGCGCGCGCGCGCGCGCGCGCGCGCGCG 833

RESULT 5  
CG376770/c 730 bp DNA linear GSS 26-AUG-2003  
LOCUS CG3CN22TH ZM\_0.7\_1.5\_KB Zea mays genomic clone ZMMBMA0776C20,  
DEFINITION genomic survey sequence.

ACCESSION CG376770  
VERSION CG376770.1 GI:34294037  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD  
AUTHORS 1. (bases 1 to 730)  
White, A., Quackenbush, J., Van Aken, S., Utterback, T.,  
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,  
Citek, R.W., Nunberg, A., Robbins, D., and Lakey, N.

TITLE Consortium for Maize Genomics  
JOURNAL Unpublished (2002)  
COMMENT Other GSSs: CG3CN22TV  
Contact: Cathy Whitelaw

TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: TR  
Class: sheared ends.

FEATURES Location/Qualifiers  
1. .730  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"

/clone="ZMMBMA0776C20"  
/clone\_lib="ZM\_0.7\_1.5\_KB"  
/note="Vector: pBSK-; Site 1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

ORIGIN  
Alignment Scores:  
Pred. No.: 2,3e-81 Length: 730  
Score: 1204.00 Matches: 237  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 41.37% Indels: 0  
DB: 29 Gaps: 0

US-10-030-884-14 (1-573) x CG376770 (1-730)

QY 121 IleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeuMetValGlnVal 140  
Db 729 ATCCCTGCTCGAGGAGCATGTACGCGCGTGTGCGCGCGCGCGCGCGCGCGCGCGCG 670  
QY 141 ValValLeuGlnCysIleIleTyrThrLeuMetLeuPheLeuPheGluTyrArgAla 160  
Db 669 GTGCTCTCCAGTGCATCATCTGTGTACAGCTCATGCTCTCTCTCTCTCTCTCTCTCT 610  
QY 161 AlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSerIleValSerPhe 180  
Db 609 GCGCGCGCGCTGCTCTGTCGACAGTTCGCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 550  
QY 181 ArgValAspSerAspValValSerLeuAlaArgGlyAspValGluLeuGluAlaGluPro 200  
Db 549 CGCGTGCAGTCCGAGCTGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 490  
QY 201 AspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGlyAspAlaGlyArgValArg 220  
Db 489 GACGCGCTCGCGCGCGCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 430  
QY 221 ValThrValArgLysSerThrSerArgSerGluAlaAlaCysSerHisSerHisSer 240  
Db 429 GTACCGGTGCGCAAGTCCACAGCTCGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCG 370  
QY 241 GlnThrMetGlnProArgValSerAsnLeuSerGlyValGluLeuTyrSerLeuGlnSer 260  
Db 369 CAGACATGACGCGCGCGCGCGCGCTGTCACACCTCTCTCTCTCTCTCTCTCTCTCTCT 310  
QY 261 SerArgAsnProThrProArgGlySerSerPheAsnHisAlaAspPheAsnIleVal 280  
Db 309 TCGCGCAACCCACCGCGCGCGCTCCAGCTTCAACACCGCGCGCGCGCGCGCGCGCGCG 250  
QY 281 GlyAlaAlaAlaLysGlyGlyGlyAlaAlaGlyAspGluGluLysGlyAlaCysGly 300  
Db 249 GCG 190  
QY 301 GlyGlyGlyGlyHisSerProGlnAlaValAlaValAlaValProAlaLysArgLys 320  
Db 189 GCG 130  
QY 321 AspLeuHisMetLeuValTrpSerSerAlaSerProValSerGluArgAlaAlaVal 340  
Db 129 GACCTGCACATGCTGCTGTGAGCTTCAGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 70  
QY 341 HisValPheGlyAlaGlyAlaAspHisAlaAspValLeuAlaLysGly 357  
Db 69 CACGCTTTCGCGCGCGCGCGCGCTGACCATGCGCGCGCGCGCGCGCGCGCGCGCGCG 19

RESULT 6  
CG317974 776 bp DNA linear GSS 26-AUG-2003  
LOCUS CGWGA91TV ZM\_0.7\_1.5\_KB Zea mays genomic clone ZMMBMA0570013,  
DEFINITION genomic survey sequence.

ACCESSION CG317974  
VERSION CG317974.1 GI:34235240  
KEYWORDS GSS.  
SOURCE Zea mays



```

ORGANISM      Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
AUTHORS      Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
              Reenick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
              Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
              Consortium for Maize Genomics
              Unpublished (2002)
              Other GSSs: OGWGA91TH
              Contact: Cathy Whitelaw
              TIGR
              9712 Medical Center Drive, Rockville, MD 20850, USA
              Tel: 301-838-5843
              Fax: 301-838-0208
              Email: whitelaw@tigr.org
              Seq primer: TF
              Class: sheared ends.
FEATURES
source
1..776
   /organism="Zea mays"
   /mol_type="genomic DNA"
   /strain="B73"
   /db_xref="taxon:4577"
   /clone="ZMMBMA0570013"
   /clone_lib="ZM.0.7.1.5_KB"
   /note="Vector: pBCKS-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Alignment Scores:
Pred. No.:      2,01e-73      Length:      776
Score:          1100.00      Matches:     225
Percent Similarity: 86.87%      Conservative: 0
Best Local Similarity: 86.87%      Mismatches:  1
Query Match:    37.80%      Indels:      34
DB:              29      Gaps:          1

US-10-030-884-14 (1-573) x CG317974 (1-776)
QY      209 ValSerSerArgGlyGlyAspAlaGlyArgValArgValThrValArgLysSerThrSer 228
Db      1  GTCTCTCCGCGCGGGACGCGCGGGCGGGTGGCGTCCAGTCCAGTCCACACG 60
QY      229 SerArgSerGluAlaAlaCysSerHisSerHisSerGlnThrMetGlnProArgValSer 248
Db      61  TCGCGCTCCGAGCGCGTGTCTGCACCTCGCACTCCAGACCATGCAGCCCGCTGTGCC 120
QY      249 AsnLeuSerGlyValGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 268
Db      121 AACCTTCGCGCGGTGGAGATCTACTCGTCGAGTGTCTCGCGCAACCCACCCCGCGCGGG 180
QY      269 SerSerPheAsnHisAlaAspPhePheAsnIleValGlyAlaAlaAlaLysGlyGly 288
Db      181 TCCAC-TTCAACCAACCGGCTCTTCAACATCGTCGCGCGCGCGCGCAAGGAGGCGGA 239
QY      289 GlyAlaAlaGlyAspGluGlyGlyAlaCysGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 308
Db      240 GGAGCGCGCGGGACAGAGAGAGAGCGCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 299
QY      309 GlnProGlnAlaValAlaValProAlaLysArgLysAspLeuHisMetLeuValTrpSer 328
Db      300 CAGCGCAGGCGCTCGCGCGTCCGCGCGCAAGAGAGGACCTGCACATGCTCTCTCGGAGC 359
QY      329 SerSerAlaSerProValSerGluArgAlaAlaValHisValPheGlyAlaGlyGlyAla 348
Db      360 TCCAGCGCTTCGCGCGGTTCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 419
QY      349 AspHisAlaAspValLeuAla-Lys-----
Db      420 GACCATGCCGACGCTCTCTCGCCCAAGGTTGGTAACAATGCGTGATCTTTTGTGTTGATTTTC 479

```

```

QY      356 ----- 356
Db      480 CTACCAAAATGTTGCTCTGCAATTTTATTGTACCAATGGTGTGCACACACAAGC 539
QY      357 -GlyAlaGlnAlaTyArgpGluTyArgpAspTyArgpSerSerArgThrLysAsnG1 376
Db      540 AGGAGGCCAGGCTACGAGAGTACGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 599
QY      376 YSerGlyGlyAlaAspLysGlyGlyProThrLeuSerLysLeuGlySerAsnSerThrAl 396
Db      600 GAGCGCGCGCGGACAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 659
QY      396 aGlnLeuTyProLysAspAspGlyGlyArgAlaAlaAlaAlaAlaAlaAlaAlaMetProAl 416
Db      660 GCAGCTGTATCCCAAGGACGAGCGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCG 719
QY      416 aSerValMetThrArgLeuLeuLeuLeuMetValTrpArgLysLeuLeuLeuLeuLeu 434
Db      720 GAGCGTATGACCGGCTATCTCTCATCATCTGTGTGGAGGAGCTGATCGGAGAC 774

RESULT 7
BU098540      658 bp      mRNA      linear      EST 29-AUG-2002
LOCUS      946136B05.v1 946 - tassell primordium prepared by Schmidt lab Zea
DEFINITION      mays cDNA, mRNA sequence.
ACCESSION      BU098540
VERSION      BU098540.1 GI:22546229
KEYWORDS      EST.
SOURCE      Zea mays
ORGANISM      Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE      1 (bases 1 to 658)
AUTHORS      Walbot,V.
TITLE      Maize ESTs from various cDNA libraries sequenced at Stanford
UNIVERSITY      University
JOURNAL      Unpublished (1999)
COMMENT      Contact: Walbot V
              Department of Biological Sciences
              Stanford University
              855 California Ave, Palo Alto, CA 94304, USA
              Tel: 650 723 2227
              Fax: 650 725 8221
              Email: walbot@stanford.edu
              Plate: 946136 row: B column: 05.
FEATURES
source
1..658
   /organism="Zea mays"
   /mol_type="mRNA"
   /cultivar="OH43"
   /db_xref="taxon:4577"
   /tissue_type="tassels"
   /dev_stage="just after the transition from vegetative to
inflorescence development"
   /lab_host="XLOLR"
   /clone_lib="946 - tassell primordium prepared by Schmidt
lab"
   /note="Organ: tassels; Vector: HybriZAP; Site 1: EcoRI;
Site 2: XhoI; George Chuck dissected immature tassels
between 1mm and 3mm. Sharon Stanfield prepared the cDNA
library in HybriZAP. Sample insert size range was 350 bp
to 3 kb with a 1 kb average."
ORIGIN
Alignment Scores:
Pred. No.:      3.08e-69      Length:      658
Score:          1043.50      Matches:     214
Percent Similarity: 98.17%      Conservative: 0
Best Local Similarity: 98.17%      Mismatches:  3
Query Match:    35.86%      Indels:      2
DB:              13      Gaps:          1

```

US-10-030-884-14 (1-573) x BU037419 (1-658)

```

Qy 323 HisMetLeuValTrpSerSerAlaSerProValSerGluArgAlaValHisVal 342
Db 7 CACATGCTCGTGGAGCTC-AGCGCTCGCCGCTCCGAGCGCGCGGTCACGTC 65
Qy 343 PheGlyAlaGlyAlaAspHisAlaAspValLeuAlaValGlyAlaGlnAlaTrpAsp 362
Db 66 TTGCGCGCCGCGCGCTGACCATGCGAGCTCTCGCAAGAGAGCCAGGCTACGAC 125
Qy 363 GluTrpGlyArgAspAspTrpSerSerArgTrpHisAsnGlySerGlyAlaAspLys 382
Db 126 GAGTACGGCGCGACGACTACAGCAGCAGCAGCAAGAGAACCGGAGCGCGCGGCAAG 185
Qy 383 GlyGlyProThrLeuSerLysLeuGlySerAsnSerThrAlaGlnLeuTrpLysAsp 402
Db 186 GGGGGCGCGAGCTGTCGAGCTGGGTCCAACTCGACGCGCGAGCTGTACCCCAAGGAC 245
Qy 403 AspGlyGluGlyArgAlaAlaAlaValAlaMetProProAlaSerValMetThrArgLeu 422
Db 246 GACGGCAGGGGAGGCGCGCGG---GCCATGCGCGCGCGAGCGTGATGACGCGGCTC 302
Qy 423 IleLeuIleMetValTrpArgLysLeuIleArgAsnProAsnThrTrpSerSerLeuIle 442
Db 303 ATCTCATCATGCTGTGTGGAGAACTGATCCGGAACCCCAACACTTACTCCAGCCTCATC 362
Qy 443 GlyValValTrpSerLeuValSerTrpArgTrpGlyIleGluMetProAlaIleIleAla 462
Db 363 GGGCTGCTGTGTCCTCGGTATCTACAGTGGGGCATCGAGATCGCGCATCGCGC 422
Qy 463 ArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeuGlyLeu 482
Db 423 CGGTGCAATTCGATCTGTGCGAGCGCGGTCTCGGATGCGCATGTTCAGCCTAGCCTG 482
Qy 483 PheMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAlaIleAlaMet 502
Db 483 TTATGCGCTGCGACCGAGATCATCGCTGCGGAAAGCTGGCGGCATCGCGATG 542
Qy 503 GlyValArgPheValAlaGlyProAlaValMetAlaAlaAlaSerIleAlaValGlyLeu 522
Db 543 GGGCTCGGTTCGTGCGAGCGCGCGGTCTATGGCGCGCGCTCCATCGCGCTGCTG 602
Qy 523 ArgGlyValLeuLeuHisIleAlaIleValGlnAlaAlaLeuProGlnGlyIle 540
Db 603 CGCGCGCTCTCTCCACATCGCATCGCTCCAGGCTGCTCGCTCAGGGATC 656

```

RESULT 8  
BU037419 627 bp mRNA linear EST 23-AUG-2002  
LOCUS 946138F12.y1 946 - tassal primordium prepared by Schmidt lab Zea  
DEFINITION mays cDNA, mRNA sequence.

ACCESSION BU037419  
VERSION BU037419.1 GI:22472939  
KEYWORDS EST.

SOURCE Zea mays

ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 627)

AUTHORS Walbot, V.

TITLE Maize ESTs from various cDNA libraries sequenced at Stanford

JOURNAL University

COMMENT Unpublished (1999)

Contact: Walbot V

Department of Biological Sciences

Stanford University

855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227

Fax: 650 725 8221

Email: walbot@stanford.edu

Plate: 946138 row: F column: 12.

Location/Qualifiers

FEATURES

source

```

1. 627
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="OH43"
/db_xref="taxon:4577"
/tissue_type="tassels"
/dev_stage="just after the transition from vegetative to
inflorescence development"
/lab_host="XLOLR"
/clone_lib="946 - tassal primordium prepared by Schmidt
lab"
/Note="Organ: tassels; Vector: HybriZAP; Site_1: EcoRI;
Site_2: XhoI; George Chuck dissected immature tassels
between 1mm and 3mm. Sharon Stanfield prepared the cDNA
library in HybriZAP. Sample insert size range was 350 bp
to 3 Kb with a 1 Kb average."

```

# ORIGIN

Alignment Scores:  
Pred. No.: 1,31e-66 Length: 627  
Score: 1008.50 Matches: 200  
Percent Similarity: 97.10% Conservative: 1  
Best Local Similarity: 96.62% Mismatches: 3  
Query Match: 34.66% Indels: 3  
DB: 13 Gaps: 2

US-10-030-884-14 (1-573) x BU037419 (1-627)

```

Qy 274 AlaAspPhePheAsnIleValGlyAlaAlaLysGlyGlyAlaAlaGlyAsp 293
Db 15 GCGGACTTCTTCAACATCGTCTTTCCGCCCCCAAGGAGGCGGAGGCGCGGGAC 74
Qy 294 GluGluLysGlyAlaCysGlyGlyGlyGlyHisSerProGlnProGlnAlaVal 313
Db 75 GAGGAGAGGGCGCATGC-----GGCGGCGAGGACACTCGCCGACGCGCGCGTC 128
Qy 314 AlaValProAlaLysArgLysAspLeuHisMetLeuValTrpSerSerAlaSerPro 333
Db 129 GCGGTGCGCGGCAAGAGAGGAGCTTCACATGCTCGTCTGGAGCTCCAGCGCTCGCCC 188
Qy 334 ValSerGluArgAlaAlaValHisValPheGlyAlaGlyGlyAlaAspHisAlaAspVal 353
Db 189 GTCTCCGAGCGCGCGCGCTGACAGTCTTCGCGCGCGCGCGCTGACCATGCGGACGTC 248
Qy 354 LeuAlaLysGlyAlaGlnAlaTyAspGluTyArgAspAspTySerSerArgThr 373
Db 249 CTCGCCAAGAGGAGCCAGCCCTACGACGATACGCGCGCGCGCGCTACAGCAGCAGGACG 308
Qy 374 LysAsnGlySerGlyGlyAlaAspLysGlyGlyProThrLeuSerLysLeuGlySerAsn 393
Db 309 AAGAACGGAGCGCGCGCGCGGACGAGGCGCGCGCGCTGTGAGCTGGGGTCCAC 368
Qy 394 SerThrAlaGlnLeuTyProLysAspAspGlyGluGlyArgAlaAlaAlaValAlaMet 413
Db 369 TCGACGCGCGAGCTGTACCCCAAGACACGCGCGAGGAGGCGCGCGCG---GCGATG 425
Qy 414 ProProAlaSerValMetThrArgLeuIleLeuIleMetValTrpArgLysLeuIleArg 433
Db 426 CGCGCGCGGAGCGGTGATGACGCGCTCATCTCATCATGTGTGAGGAAGCTGATCCGG 485
Qy 434 AsnProAsnThrTySerSerLeuIleGlyValValTrpSerLeuValSerTyArgTrp 453
Db 486 AACCACCAACACTTACTCCAGGCTCATCGCGCTCGTCTGGTCCCTGGTATCTCTACAGGTGG 545
Qy 454 GlyIleGluMetProAlaIleAlaArgSerIleSerIleLeuSerAspAlaGlyLeu 473
Db 546 GGCATCGAGATGCCAGCATCATGCCCGGTGATTCGATTCCTGCTCGGACGCGGTCTC 605
Qy 474 GlyMetAlaMetPheSerLeu 480
Db 606 GGGATGGCCATGTTTCAGCCTA 626

```

RESULT 9

AY106004		AY106004	1426 bp	mRNA	linear	HTC 16-OCT-2002
LOCUS	Zea mays	P0072074	mRNA sequence.			
DEFINITION	Zea mays	AY106004				
ACCESSION	Zea mays	AY106004				
VERSION	Zea mays	AY106004.1	GI:21209082			
KEYWORDS	HTC.					
SOURCE	Zea mays					
ORGANISM	Zea mays					
Eukaryota; Viridiplantae;						
Streptophyta; Embryophyta;						
Tracheophyta; Magnoliopsida;						
Liliopsida; Poales; PACCAD						
clade; Panicoideae; Andropogoneae; Zea.						
1 (bases 1 to 1426)						
Hanev,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whiteitt,M.S.,						
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.						
Maze Mapping Project/DuPont Consensus Sequences for Design of						
Overgo Probes						
Unpublished (2002)						
2. (bases 1 to 1426)						
Coe,E.H.						
Direct Submission						
Submitted (23-APR-2002) Maize Mapping Project, University of						
Missouri, Columbia, MO 65211, USA						
If you are interested in getting corresponding physical clones,						
these are publicly available from ZMDB and may be found by BLAST						
searching at MSU, maizemap.org; ZMDB, www.zmdb.iastate.edu; TIGR,						
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the						
maize cDNA sequences is either Virginia Walbot, Stanford or Pat						
Schnable, Iowa State, then clones may be requested from ZMDB:						
www.zmdb.iastate.edu.						
Location/Qualifiers						
1..1426						
/organism="Zea mays"						
/mol_type="mRNA"						
/db_xref="MaizeDB:633932"						
/db_xref="taxon:4577"						
/clone_lib="Maize Mapping Project/DuPont Consensus						
Library"						
/note="this sequence is part of a project of EST						
assemblies resulting from the application of public						
configs to seed DuPont contigs; this resource was						
assembled by DuPont as part of a collaboration for the						
overgo addressing of BACs in conjunction with the Maize						
Mapping Project"						
Alignment Scores:						
Pred. No.: 9,83e-66 Length: 1426						
Score: 1004.00 Matches: 222						
Percent Similarity: 64.89% Conservative: 33						
Best local Similarity: 56.49% Mismatches: 68						
Query Match: 34.50% Indels: 70						
DB: 11 Gaps: 12						
US-10-030-884-14 (1-573) x AY106004 (1-1426)						
QY 1 MetTleThrAlaLeuAspLeuTyTHisValLeuThrAlaValProLeuTyValala 20						
320 ATGATTACGGGCAGCGACTTTCCACCATCATGCCCGGTGGTGCGGTGTACGTGGCG 379						
QY 21 MetThrLeuAlaTygGlySerValArgTriPTrpArgIlePherThrProAspGlnCysser 40						
380 ATGATCTTGGCTACGGGTTCGGTTCGGTGGTGGCGCATCTTCTCGCGCACCAAGTCCTCC 439						
QY 41 GlyIleasnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleaser 60						
440 GGAGTAAACCGGCTTCGTGGCGCTCTTCGGGTTCGGTTCATCGCGCGCACGAGTCACATCCC 499						
QY 61 ThrAsnAspPropPheAlaMetAsnLeuArgPheLeuAlaAalaAspThrLeuGlnIysVal 80						
500 ACCAACACCCCTACCACTGAACCTGGCGTTTCATCGCGCGCACGAGTCGAGAAGTC 559						
QY 81 AlaValLeuAlaLeuAla-----LeuAlaserArgGlyLeuSerSerPro 96						

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoidae; Andropogoneae; Zea.  
 1 (bases 1 to 610)  
 Walbot, V.  
 Maize ESTs from various cDNA libraries sequenced at Stanford  
 University  
 Unpublished (1999)  
 Contact: Walbot V  
 Department of Biological Sciences  
 Stanford University  
 855 California Ave, Palo Alto, CA 94304, USA  
 Tel: 650 723 2227  
 Fax: 650 725 8221  
 Email: walbot@stanford.edu  
 Plate: 946150 row: G column: 03.  
 Location/Qualifiers  
 1. 610  
 /organism="Zea mays"  
 /mol\_type="mRNA"  
 /culti\_var="OH43"  
 /db\_xref="taxon:4577"  
 /tissue\_type="tassels"  
 /dev\_stage="just after the transition from vegetative to  
 inflorescence development"  
 /lab\_host="XJOLR"  
 /clone\_lib="946 - tassels primordium prepared by Schmidt  
 lab"  
 /note="Organ: tassels; Vector: HybriZAP; Site 1: EcoRI;  
 Site 2: XhoI; George Chuck dissected immature tassels  
 between 1mm and 3mm. Sharon Stanfield prepared the cDNA  
 library in HybriZAP. Sample insert size range was 350 bp  
 to 3 Kb with a 1 Kb average."

ORIGIN

Alignment Scores:  
 Pred. No.: 6,05e-66 Length: 610  
 Score: 999.50 Matches: 197  
 Percent Similarity: 98.01% Conservative: 0  
 Best Local Similarity: 98.01% Mismatches: 3  
 Query Match: 34.35% Indels: 1  
 DB: 13 Gaps: 1

US-10-030-884-14 (1-573) x BU080190 (1-610)

QY 303 GlyGlyGlyHisSerProGlnProGlnAlaValAlaValProAlaLysArgLysAspLeu 322  
 Db 3 GCGGAGGACACTCGCCGAGCGCGCTGCGCGTGCCTGCGGCAAGAGGAGGACCTG 62  
 QY 323 HisMetLeuValTrpSerSerAlaSerProValSerGluArgAlaAlaValHisVal 342  
 Db 63 CACATGCTCGTGTGAGCTCCAGCGCTCGCCCTCTCCGAGCGCGCGCTGACGTC 122  
 QY 343 PheGlyAlaGlyGlyAlaAspHisAlaAspValLeuAlaLysGlyAlaGlnAlaTyAsp 362  
 Db 123 TTGCGCGCGCGCGCTGACCTGCGGACGCTCTCGCCAAAGAGAGCCAGGCTACGAC 182  
 QY 363 GluTyGlyArgAspAspTySerSerArgThrIysAsnGlySerGlyAlaAspLys 382  
 Db 183 GAGTACGCGCGCGAGCTACAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 242  
 QY 383 GlyGlyProThrLeuSerLysLeuGlySerAsnSerThrAlaGlnLeuTyProLysAsp 402  
 Db 243 GCGGCGCGCGAGCTGTGAGCTGGGTCCTACTCGACGCGCGAGCTGTACCCCAAGGAC 302  
 QY 403 AspGlyGluGlyArgAlaAlaValAlaMetProAlaSerValMetThrArgLeu 422  
 Db 303 GACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 359  
 QY 423 IleLeuIleMetValTrpArgLysLeuIleArgAnProAsnThrTySerSerLeuIle 442  
 Db 360 ATCCCTCATCTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 419  
 QY 443 GlyValValTrpSerLeuValSerTyArgTrpGlyIleGluMetProAlaIleAla 462

Db 420 GCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 479  
 QY 463 ArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeuGlyLeu 482  
 Db 480 CGGTGATTCGATCTGTCGACGCGGCTCTCGGATGGCCATGTCAGCCTAGGCGCTG 539  
 QY 483 PheMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAlaAlaIleAlaMet 502  
 Db 540 TTCATGGGCGCTCAGCCGAGGATCATCGTGGGGAACAAGCTGCGGGCATCGCGATG 599  
 QY 503 Gly 503  
 Db 600 GGG 602

RESULT 11  
 LOCUS CA831110 604 bp mRNA linear EST 12-DEC-2002  
 DEFINITION 1117015F07.y1 1117 - Unigene V from Maize Genome Project Zea mays  
 CDNA, mRNA sequence.  
 ACCESSION CA831110 GI:26558875  
 VERSION CA831110  
 KEYWORDS EST.  
 SOURCE Zea mays  
 ORGANISM Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoidae; Andropogoneae; Zea.  
 1 (bases 1 to 604)  
 Walbot V.  
 Maize ESTs from various cDNA libraries sequenced at Stanford  
 University  
 Unpublished (1999)  
 Contact: Walbot V  
 Department of Biological Sciences  
 Stanford University  
 855 California Ave, Palo Alto, CA 94304, USA  
 Tel: 650 723 2227  
 Fax: 650 725 8221  
 Email: walbot@stanford.edu  
 Plate: 1117015 row: F column: 07.  
 Location/Qualifiers  
 1. 604  
 /organism="Zea mays"  
 /mol\_type="mRNA"  
 /db\_xref="dbEST:946138F12.y1"  
 /clone\_lib="1117 - Unigene V from Maize Genome Project"  
 /note="This library represents the unique genes found in  
 the fifth round of EST sequencing at Stanford University  
 for the maize genome project. Sequences are present from  
 library 946. Contigs were assembled using ZmBAssembler  
 and 2 representatives from each contig were selected for  
 the Unigene set. All singlets were also selected."

ORIGIN

Alignment Scores:  
 Pred. No.: 8,45e-66 Length: 604  
 Score: 997.50 Matches: 197  
 Percent Similarity: 97.52% Conservative: 0  
 Best Local Similarity: 97.52% Mismatches: 2  
 Query Match: 34.28% Indels: 3  
 DB: 14 Gaps: 2

US-10-030-884-14 (1-573) x CA831110 (1-604)

QY 274 AlaAspPhePheAsnIleValGlyAlaAlaAlaLysGlyGlyAlaAlaGlyAsp 293  
 Db 6 GCGGACTTCTTCAACATCGTCTGCTGCGGCGGCAAGGAGGAGGAGGAGGAGGAGGAG 65  
 QY 294 GluGluLysGlyAlaCysGlyGlyGlyGlyGlyHisSerProGlnProGlnAlaVal 313  
 Db 66 GAGGAGAGGGGCGCATGC-----GGCGGCGGAGGAGACTTCGCGCGCAGCGGCGGTC 119

library 946. Contigs were assembled using ZmDBAssembler and 2 representatives from each contig were selected for the Unigene set. All singlets were also selected."

ORIGIN

Alignment Scores:  
Pred. No.: 1.14e-64 Length: 630  
Score: 983.00 Matches: 202  
Percent Similarity: 96.19% Conservative: 0  
Best Local Similarity: 96.19% Mismatches: 5  
Query Match: 33.78% Indels: 3  
DB: 14 Gaps: 2

US-10-030-884-14 (1-573) x CA830783 (1-630)

QY 323 HisMetLeuValTrrSerSerAlaSerProValSerGlu-ArGAlaAlaValHisVa 342  
Db 7 CACATGCTCGTGTGGAGCTCCCTTTCTCGCCGCTCCGAGCGCGC---CGCGTGCACGT 63

QY 342 lPheGlyAlaGlyGlyAlaAspHisAlaAspValLeuAlaLysGlyAlaGlnAlaTyrHis 362  
Db 64 CTTTCGCGCGCGCGCGCTGACCATGCCAGCTCTCCCAAGAGAGCCGAGCCCTACCA 123

QY 362 pGluTyrGlyArGAspAspTyr-SerSerArgThrLysAsnGlySerGlyGlyAlaAspLys 382  
Db 124 CAGATACGGCGCGCGAGCACTACAGCAGCAGAGCAGAGAACGCGGCGCGCGGACAA 183

QY 382 sGlyGlyProThrLeuSerLysLeuGlySerLysSerThrAlaGlnLeuTyrProLysHis 402  
Db 184 GGGCGCGCGCGAGCTGCTCGAAGCTGGGGTCCAACTCGACGCGCGCAGCTGTATCCCAAGGA 243

QY 402 pAspGlyGlyGlyArGAlaAlaValAlaMetProProAlaSerValMetThrArgLe 422  
Db 244 CGACGCGGAGGAGGAGCAGCGCGCG---CGATGCGCGCGCGCGGCTGATCAGCGGCT 300

QY 422 uilleLeuileMetValTrrArgLysLeuileArgAsnProAsnThrTyrSerSerLeuile 442  
Db 301 CATCTCATCATGCTGTGGAGGAAGCTGATCCGGAACCCCAACACTTACTCCAGCCTCAT 360

QY 442 eGlyValValTrrSerLeuValSerTyrArgThrGlyLysLeuMetProAlaLysLeAl 462  
Db 361 CGCGCTGCTGTGCTCCCTGCTATCTTACAGGTGGGCACTCGAGATCCAGCGCATCATCGC 420

QY 462 aArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeuGlyLe 482  
Db 421 CGGTCGATTTGATCTCTGTCGAGCGGCTCTCGGATGGCCATGTTCCAGCCTAGGCT 480

QY 482 uPheMetAlaLeuGlnProArgIleleAlaCysGlyAsnLysLeuAlaLysLeAlaMe 502  
Db 481 GTTCATGGCGCTCGACCGCAGGATCATCGCTGCGGAACAAGCTGCGGCCCATCCGAT 540

QY 502 tGlyValArgPheValAlaGlyProAlaValMetAlaAlaAlaSerIleAlaValGlyLe 522  
Db 541 GGGGTCGCTGCTGTCGAGGCGCGCGCTGATGGCGCGCGCTCATCGCGCTCGCTCT 600

QY 522 uArgGlyValLeuLeuHisIleAlaile 531  
Db 601 GCGCGCGCTCTCTCCACATCGCCATC 628

RESULT 13  
CG358058 726 bp DNA linear GSS 26-AUG-2003  
LOCUS CG358058  
DEFINITION CG358058 ZM 0.7\_1.5\_KB Zea mays genomic clone ZM58Ma0724C07,  
genomic survey sequence.  
ACCESSION CG358058  
VERSION CG358058.1 GI:34275325  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 726)

314 AlaValProAlaLysArGAspLeuHisMetLeuValTrrSerSerSerAlaSerPro 333  
Db 120 GCCGTGCGCGCGCGCGAGGAGCACTGCACATGCTCTGCTGAGCTCAGCGCTCGCCC 179

QY 334 ValSerGluArGAlaAlaValHisValPheGlyAlaGlyGlyAlaAspHisAlaAspVal 353  
Db 180 GTCTCCGAGCGCGCGCGCGTGCAGCTCTTCGCGCGCGCGCGCTGACCACTGCGACGTC 239

QY 354 LeuAlaLysGlyAlaGlnAlaTyrAspGlyTyrGlyArGAspAspTyrSerSerArgThr 373  
Db 240 CTCGCAAGAGGAGCCAGCCCTACGACGAGTACGCGCGCGAGCACTACAGCAGCAGGACG 299

QY 374 LysAsnGlySerGlyGlyAlaAspLysGlyGlyProThrLeuSerLysLeuGlySerAsn 393  
Db 300 AAGAACGGGAGCGCGCGCGAGCAAGCGCGCGCGCTGTCGAGCTGGGTCCCAAC 359

QY 394 SerThrAlaGlnLeuTyrProLysAspAspGlyGlyGlyArgAlaAlaAlaValAlaMet 413  
Db 360 TCGACGGCGCGAGCTGTATCCCAAGAGCAGCGCGAGGAGGCGCGCGCG---GCCATG 416

QY 414 ProProAlaSerValMetThrArgLeuileLeuileMetValTrrArgLysLeuileArg 433  
Db 417 CGCGCGCGGAGGCTGATGACGCGGCTCATCTCATGCTGTGGAGGAGCTGATCCGG 476

QY 434 AnProAsnThrTyrSerSerLeuileGlyValValTrrSerLeuValSerTyrArgThr 453  
Db 477 AACCCCAACACTTACTCCAGCCTCATCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 536

QY 454 GlyLysGlyMetProAlaLysLeuAlaArgSerIleLeuSerAspAlaGlyLeu 473  
Db 537 GGCATCGAGATCCAGCGCATATCCCGCGTTCGATTCGATTCCTGTCGAGCGCGGCTCTC 596

QY 474 GlyMet 475  
Db 597 GGGATG 602

RESULT 12  
CA830783 630 bp mRNA linear EST 12-DEC-2002  
LOCUS 1117011D11.y1 1117 - Unigene V from Maize Genome Project Zea mays  
DEFINITION cDNA, mRNA sequence.  
ACCESSION CA830783  
VERSION CA830783.1 GI:26558548  
KEYWORDS EST.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 630)  
AUTHORS Walbot V.  
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford  
UNIVERSITY University  
CONTACT Walbot V  
DEPARTMENT Department of Biological Sciences  
STANDFORD STANFORD UNIVERSITY  
855 CALIFORNIA AVE, PALO ALTO, CA 94304, USA  
TEL: 650 723 2227  
FAX: 650 725 8221  
EMAIL: walbot@stanford.edu  
PLATE: 1117011 row: D column: 11.  
LOCATION/Qualifiers  
1. .630  
/organism="Zea mays"  
/mol\_type="mRNA"  
/db\_xref="taxon:4577"  
/clone\_lib="1117 - Unigene V from Maize Genome Project"  
/notes="This library represents the unique genes found in  
the fifth round of EST sequencing at Stanford University  
for the maize genome project. Sequences are present from

AUTHORS Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Renick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nurnberg, A., Robbins, D., and Lakey, N.

TITLE Consortium for Maize Genomics

JOURNAL Unpublished (2002)

COMMENT Other\_GSSs: OGI16K16TV  
Contact: Cathy Whitelaw

TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: TR  
Class: sheared ends.

FEATURES source  
1. 726  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clones="ZMMBMA0724C07"  
/clone\_lib="ZM 0.7 1.5 KB"  
/note="Vector: pBSK-; Site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"

ORIGIN  
Alignment Scores:  
Pred. No.: 5,54e-63 Length: 726  
Score: 962.00 Matches: 192  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 33.06% Indels: 0  
DB: 29 Gaps: 0

US-10-030-884-14 (1-573) x CG358058 (1-726)

Qy 1 MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValProLeuTyrValAla 20  
Db 150 ATGATCACCGCGCTGGACCTTACACGCTGTGACGGGGTGGCGCTGTACGTGCC 209  
Qy 21 MetThrLeuAlaTyrGlySerValArgTyrTrpArgIlePheThrProAspGlnCysSer 40  
Db 210 ATGAGCTGGCTAGCGCTCGCTGGCTGGTGGCGATCTTACCGCGGACGAGTGCTCC 269  
Qy 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
Db 270 GGGATCAACCGCTTCTGGCGCTCTTGGCGCTGGCTGGCTCTCTCTCTCCATCTCC 329  
Qy 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80  
Db 330 ACCAACGACCCCTTCGCCATGAACCTGGCTTCTTGGCGCGGACACGCTGCAGAGGTG 389  
Qy 81 AlaValLeuAlaLeuLeuAlaLeuAlaSerArgGlyLeuSerSerProArgAlaLeuGly 100  
Db 390 GCCGTCTTCCGGCTGCTGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 449  
Qy 101 LeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThrLeuValMetGly 120  
Db 450 CTCGACTGGAGCATCACGCTCTTCTCCCTCTCCAGCTTCCCAACAGCTGCTCATGGGC 509  
Qy 121 IleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeuMetValGlnVal 140  
Db 510 ATCCCGCTGCTGGAGGCGATGTACGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 569  
Qy 141 ValValLeuGlnCysIleIleTyrThrLeuMetLeuPheLeuPheGluTyrArgAla 160  
Db 570 GTCGTCTTCCAGTGCATCATCTGGTACACGCTTCAGCTCTTCTCTCTTCAGTACCGGCC 629  
Qy 161 AlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSerIleValSerPhe 180  
Db 630 GGGCGCGCGCTCTCTCGACGAGTTCCTCCGACGGCGCGCGCTCCATGCTCTCTCTC 689  
Qy 181 ArgValAspSerAspValValSerLeuAlaArgGly 192

Db 690 CGCGTCGACTCCGACGTCGTCTCGTCCGCCAGGGGG 725

RESULT 14  
AY110024  
LOCUS Zea mays CL464\_3 mRNA sequence. 1221 bp linear HTC 17-OCT-2002  
DEFINITION Zea mays CL464\_3 mRNA sequence.  
ACCESSION AY110024  
VERSION AY110024.1 GI:21214085  
KEYWORDS HTC.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 1221)  
AUTHORS Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitesitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V. Design of Maize Mapping Project/DuPont Consensus Sequences for the Overgo Probes  
JOURNAL Unpublished (2002)  
REFERENCE 2 (bases 1 to 1221)  
AUTHORS Coe, E.H.  
TITLE Direct Submission  
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA  
COMMENT If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizegap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.

FEATURES Location/Qualifiers  
1. 1221  
/organism="Zea mays"  
/mol\_type="mRNA"  
/db\_xref="MaizeDB:632107"  
/db\_xref="taxon:4577"  
/clone\_lib="Maize Mapping Project/DuPont Consensus Library"  
/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

ORIGIN  
Alignment Scores:  
Pred. No.: 1.14e-60 Length: 1221  
Score: 936.00 Matches: 212  
Percent Similarity: 63.75% Conservative: 35  
Best Local Similarity: 54.50% Mismatches: 75  
Query Match: 32.16% Indels: 66  
DB: 11 Gaps: 12

US-10-030-884-14 (1-573) x AY110024 (1-1221)

Qy 1 MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValProLeuTyrValAla 20  
Db 142 ATGATCACCGCGACGACTTACACGCTGTACCGCTGTCCGTGTACGTGCC 201  
Qy 21 MetThrLeuAlaTyrGlySerValArgTyrTrpArgIlePheThrProAspGlnCysSer 40  
Db 202 ATGATCTCGGTACGATCCGTCAGTGGTGGCGATCTTACCGCGGACGAGTCTCG 261  
Qy 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
Db 262 GGGATCAACCGCTTCTGGCGCTCTTCCGCGCGCTCTCTCTCTCCATTCANNNN 321  
Qy 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80  
DB: 11 Gaps: 12

[illegible]



```
Db 484 CCAGCGATCATCGCCCGTGCATTTCGATCTCTGAGACGGGTCCTCGGATGCGCATG 543
Qy 478 PheSerLeu----- 480
Db 544 TTCAGCCTAGGTGTGTGTGTCTTACTTGTACGGGCTCCCGTGTCTGCTGACCCGCTG 603
Qy 481 -----GlyLeuPheMetAlaLeu 486
Db 604 CAATGCGAGCTCTGAACCGCGCCTTGCCCTGTGTTCAGGCTGTTCATGGGCTG 663
Qy 487 GlnProArgIleIleAlaCysGlyAsnIlyLeuAlaIleAlaMetGlyValArgPhe 506
Db 664 CAGCCGAGGATCATCGGCTGCGGAAACAAGCTGGCGCCATCGCGATGGGCGTCCGCTTC 723
Qy 507 ValAlaGlyProAlaValMetAlaAlaAlaSerIleAlaValGlyLeuArgGlyValLeu 526
Db 724 GTCGAGGCCCGCGGTTCATGCGCGCCTCCATCGCGTCTGCGGCGTCTC 783
Qy 527 LeuHisIleAlaIleVal----- 532
Db 784 CTCCACATCGCCATCGT-CCAGGTACGCGCCCTGTGCTTGAACGTAAACGGTGGAGTTGT 842
Qy 533 -----GlnAlaAlaLeuProGln 538
Db 843 GATGGAACAGTAGGACCAACATTCGTTTCCCTTCTTCTCCAGGCTGCTCTGCCCTCAG 902
Qy 539 GlyIleValProPheValPheAlaLysGluTyrGlyValHisProAspIleLeuSerThr 558
Db 903 GGGATCGTGGCGTTCGTGTTCCCAAGGAGTACGGCGTTTCATCCCGACATCCTGAGCACA 962
Qy 559 AlaTyrGly 561
Db 963 GCGTATGGT 971
```

Search completed: March 3, 2004, 12:16:03  
Job time : 4006 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 3, 2004, 09:13:09 ; Search time 116 Seconds  
(without alignments)  
2741.266 Million cell updates/sec

Title: US-10-030-884-14

Perfect score: 2910  
Sequence: 1 M1RALDLYHVTVAVPLVA.....DILSTAYGPTSGFITCHS 573

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Dgapop 6.0 , Dgapext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+p2n.model -DEV=xlh  
-Q=/cgm2\_1/USPTO\_gpool/US10030884/runat 24022004 103923 1972/app\_query.fasta\_1.711  
-DB=Issued Patents NA -QMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPEXT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -LOCAL=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10030884 @CGN 1.1 56 @runat 24022004 103923 1972 -NCPU=6 -ICPU=3  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAEXT=0.5 -FGAPOP=6  
-FGAEXT=7 -YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:

1: /cgm2\_6/ptodata/2/ina/5A.COMB.seq:\*  
2: /cgm2\_6/ptodata/2/ina/5B.COMB.seq:\*  
3: /cgm2\_6/ptodata/2/ina/6A.COMB.seq:\*  
4: /cgm2\_6/ptodata/2/ina/6B.COMB.seq:\*  
5: /cgm2\_6/ptodata/2/ina/PTCUS.COMB.seq:\*  
6: /cgm2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	172.5	5.9	963	US-09-489-039A-6633	Sequence 6633, Ap
2	152.5	5.2	68750	US-09-335-409-1	Sequence 1, Appli
3	152.5	5.2	68750	US-09-568-102-1	Sequence 1, Appli
4	152.5	5.2	68750	US-09-567-969-1	Sequence 1, Appli
5	152.5	5.2	68750	US-09-568-480-1	Sequence 1, Appli
6	152.5	5.2	68750	US-09-568-486-1	Sequence 1, Appli
7	152.5	5.2	68750	US-09-568-472-1	Sequence 1, Appli
8	152.5	5.2	68750	US-09-567-899-1	Sequence 1, Appli
9	151.5	5.2	71989	US-09-443-501A-2	Sequence 2, Appli
C 10	144	4.9	4403765	US-09-103-840A-2	Sequence 2, Appli
C 11	139	4.8	4411529	US-09-103-840A-1	Sequence 1, Appli
12	138.5	4.8	2214	US-08-864-038A-1	Sequence 1, Appli

13	138.5	4.8	3331	3	US-08-864-038A-2	Sequence 2, Appli
14	138.5	4.8	3331	3	US-08-864-038A-4	Sequence 4, Appli
15	138	4.7	1682	4	US-09-220-132-82	Sequence 82, Appli
16	138	4.7	4411529	3	US-09-103-840A-1	Sequence 1, Appli
C 17	136	4.7	3211	2	US-08-574-959A-8	Sequence 8, Appli
C 18	136	4.7	3211	3	US-09-357-014-8	Sequence 8, Appli
C 19	136	4.7	3901	2	US-08-574-959A-6	Sequence 6, Appli
C 20	136	4.7	3901	3	US-09-357-014-6	Sequence 6, Appli
C 21	135	4.6	4403765	3	US-09-103-840A-2	Sequence 2, Appli
C 22	134.5	4.6	4446	4	US-09-231-899-69	Sequence 69, Appli
C 23	134	4.6	3129	4	US-09-252-991A-13873	Sequence 13873, A
C 24	134	4.6	8211	4	US-09-252-991A-13656	Sequence 13656, A
C 25	133.5	4.6	1710	2	US-08-912-794-1	Sequence 1, Appli
C 26	132	4.5	47981	4	US-09-679-279-1	Sequence 1, Appli
C 27	131	4.5	2561	4	US-09-616-289-48	Sequence 48, Appli
C 28	131	4.5	4131	4	US-09-252-991A-13773	Sequence 13773, A
C 29	130.5	4.5	2385	4	US-09-252-991A-11998	Sequence 11998, A
C 30	129.5	4.5	3478	1	US-08-530-492-1	Sequence 1, Appli
C 31	129.5	4.5	3478	3	US-08-906-517-1	Sequence 1, Appli
C 32	128.5	4.4	1664976	4	US-08-916-421B-1	Sequence 1, Appli
C 33	128	4.4	34094	4	US-09-292-034-1	Sequence 1, Appli
C 34	127.5	4.4	1746	4	US-09-252-991A-1163	Sequence 1163, Ap
C 35	127.5	4.4	2451	4	US-09-252-991A-7534	Sequence 7534, Ap
C 36	127.5	4.4	2661	4	US-09-252-991A-7807	Sequence 7807, Ap
C 37	126.5	4.3	1768	4	US-09-833-381-523	Sequence 523, Ap
C 38	126.5	4.3	1803	1	US-08-821-119-18	Sequence 18, Appli
C 39	126.5	4.3	1803	2	US-08-821-118-1	Sequence 1, Appli
C 40	126.5	4.3	4435	4	US-09-612-402B-1	Sequence 1, Appli
C 41	126	4.3	3102	4	US-09-489-039A-6353	Sequence 6353, Ap
C 42	126	4.3	77536	4	US-09-410-551B-1	Sequence 1, Appli
C 43	125.5	4.3	2868	4	US-09-252-991A-7683	Sequence 7683, Ap
C 44	125.5	4.3	13842	3	US-09-105-537-30	Sequence 30, Appli
C 45	125.5	4.3	13987	2	US-08-804-227C-13	Sequence 13, Appli

#### ALIGNMENTS

##### RESULT 1

US-09-489-039A-6633  
; Sequence 6633, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 6633  
; LENGTH: 963  
; TYPE: DNA  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-6633

##### Alignment Scores:

Pred. No.: 7.81e-07 Length: 963  
Score: 172.50 Matches: 71  
Percent Similarity: 42.02% Conservative: 37  
Best Local Similarity: 27.63% Mismatches: 131  
Query Match: 5.93% Indels: 18  
DB: 4 Gaps: 8

US-10-030-884-14 (1-573) x US-09-489-039A-6633 (1-963)

QY 300 GYGLYGLYGLYGLYGLYHISserProGlnProGlnAlaValAlaValProAlaLysArg 319

Db 162 GGTACAAACGCGCGCGCGGATCGTCCAGTCGCGCTGGTGTGTGACCGG 221

QY 320 LysAspLeuHisMetLeuValTrpSerSerAlaSerProValSerGluArgAlaAa 339



[illegible]

```

RESULT 3
US-09-568-102-1
; Sequence 1, Application US/09568102
; Patent No. 6346404
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/568,102
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 68750
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
US-09-568-102-1

Alignment Scores:
Pred. No.: 0.047
Score: 152.50
Percent Similarity: 29.68%
Best Local Similarity: 21.12%
Query Match: 5.24%
DB: 4

Length: 68750
Matches: 106
Conservative: 43
Mismatches: 164
Indels: 190
Gaps: 14

US-10-030-884-14 (1-573) x US-09-568-102-1 (1-68750)

```

	:    :        :    :	
D <sub>b</sub>	28213 GTGCTGGCGGGGTGTTCTTCGTGGAAGATGGGGTGAGGCTCGTGGCGCGCG-	28266
Q <sub>y</sub>	138 ValGlnValValValLeuGlnCysile-leItrpyrThrMetMetPheLeuPheGlu	157
D <sub>b</sub>	28266 -----	28266
Q <sub>y</sub>	158 TyrArgAlaAraArgAlalaLeuValLeuAspGlnPheProaspGlyAlaAlaAserIle	177
D <sub>b</sub>	28267 ---CGGGCGGCTGATGCAAGGGCTCTCCGCCGGCGCGCGCATGTGTGCCTCGAGC---	28320
Q <sub>y</sub>	178 ValSerPheArgValAspSerAspValValSerLeuAlaArgGlyAspValGluLeuGlu	197
D <sub>b</sub>	28320 -----	28320
Q <sub>y</sub>	198 AlaGluProAspGlyValAlaGlyAlaGlyAlaVal-----	209
D <sub>b</sub>	28321 GCGGAGGCGGAGGTGGCGCGCGCGGTGGCCGCCACCGCGCGTGGTGCATGCGCGCGC	28380
Q <sub>y</sub>	210 -----SerSerArgGlyGlyAspAlaGlyArg-----ValArg	220
D <sub>b</sub>	28381 GGTCAATGGCGCGGACGAGTGGTATCGCGGCGTGGAGCAACGCGGTGCAGCGCATCGC	28440
Q <sub>y</sub>	221 ValThrValArgLysSerThr-SerSerArgSerGluAlaAlaCys--SerHis--SerHi	239
D <sub>b</sub>	28441 GCGGGGTTTCGGCGCGCGCGCGCGCACCAACGCGCTGCATGCTCGCACGCGTCCCA	28500
Q <sub>y</sub>	239 sSerGlnThrMetGlnPro-----ArgValSerAsnLeuSerGlyVa	253
D <sub>b</sub>	28501 CTCGCGCTGATGGAAACCAGTCTCGAGGACTTCGGGCGGGTGGCGGC-GTGCGTAGCT	28559
Q <sub>y</sub>	253 lGlulleYrSerLeuGlnInSerSerArgAsnProThrProArgGlySerSerPheAsnHi	273
D <sub>b</sub>	28560 ACCGGCGCCAAGCGTTTTCGTGTGTGAGCAAACCTGAGCGGGAAGTGGTGC- --GGACG	28616
Q <sub>y</sub>	273 sAlaAspPhePheAsnIleValGlyAlaAlaAlaLysGlyGlyGlyAlaAlaGlyAs	293
D <sub>b</sub>	28617 AGCTGAGCGCGCGGGGTACTGGGTGGCGACGTCGGGAGGCGGTGCGCTTCGCGGACG	28676
Q <sub>y</sub>	293 pGluglunLysGlyAlaCysGlyGlyGlyGlyGlyGlyHisSerProGlnProGlnAlaVa	313
D <sub>b</sub>	28677 GGGTGAAGCGCTGCACAAGACCGGTGGCGGACGCTGTCGAAGTGGGCGGACCGGA	28736
Q <sub>y</sub>	313 lAlaValProAlaLysA-GlysAspLeuHisMetLeuValTrpSerSerAlaSerPr	333
D <sub>b</sub>	28737 CGTGTCT-----	28743
Q <sub>y</sub>	333 oValSerGluArgAlaAlaValHisVal:PheGlyAlaGlyGlyAlaAspHisAlaSpVa	353
D <sub>b</sub>	28744 -----CGGGCTGTTCAGCCTCTCTCGCGGAGCGGAGCGACGCTCTCGCGCT	28793
Q <sub>y</sub>	353 lLeuAlaLysGlyAlaGlnAlaTyArgPgluTyArgAspAspTyfSerSerArgTh	373
D <sub>b</sub>	28794 CGTTCGCGC-----	28808
Q <sub>y</sub>	373 rLyAsnGlySerGlyAlaAspLysGlyGlyProThrLeuSerLysLeuGlySerAs	393
D <sub>b</sub>	28809 GCGAGGAGCTCGGGGGTGTTCGAGGCGCTGGCGAGGCTGTGGCGCGGGGTCTCGG	28868
Q <sub>y</sub>	393 nSerThrAlaGlnLeuTyProLysAspAspGlyGluGlyArgAlaAlaValAlaMe	413
D <sub>b</sub>	28869 TCAGCTGGCGGGCGCTTTCGCCACCGGTGGGCGGGTGGCGCTGCCGACCTATCCGT	28928
Q <sub>y</sub>	413 tProProAlaSerValMetThr-----	420
D <sub>b</sub>	28929 GGCAGCGGCAGCGGTACTGGCCGACATCGAGCCTGACAGCCGTGCGCACGACCGCGG	28988
Q <sub>y</sub>	421 -----ArgLeuIleLeulle-MetValTrpArgLysLeulleArgAsn----	434
D <sub>b</sub>	28989 ATCCGACCCAGGCTGGTITCATCGCTGGACTGGCCGGAGATACCTCGGAGCTCCAGA	29048
Q <sub>y</sub>	434 -----	434

118 ValMetGlyIleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeuMet 137 QY

Db 29049 AATCAGAGGAGCGAGCGCGGAGCTGGCTGTTATGGCGGATAAGGTTGAGTCGGCG 29108  
Qy 435 -----ProA 436  
Db 29109 AGCGGTCTGCTGAGCGCTGTGCACACGTGGACTTCCATGCTGCTGCTCCATGCGCGG 29168  
Qy 436 snThrTyrSerSerLeuIleGlyValVal-----Trps 447  
Db 29169 CAGAGACATCCGAGCGCGGAGCTGGTACCGAGGCTGCGCGGCGTGAAGCGATTGCG 29228  
Qy 447 erLeuValSerTyrArgTTPGlyVileGluMetProAlaIleleAlaIleAlaIleAlaIle 467  
Db 29229 AGGTAGTCTACCTGTGGGTCTGGACCGCTGCGTGGTGGTGGTGGTGGTGGTGGTGGTGG 29288  
Qy 467 leLeuSerAspAlaGly-----LeuGlyMetAlaMetPheSerL 480  
Db 29289 AGATCGGCGAGCGGACCGCTGCTGTACCGCGCGGCTGCTGGCTGGCTGGCTGGCTGGCTGG 29347  
Qy 480 euGlyLeuPheMetAlaLeuGlnProArgTlleAlaCysGlyAsnLysLeuAlaIleAlaI 500  
Db 29348 -----AGCACCGTGTCTTGTTCGCCCGGACTCTGGGTGCG 29381  
Qy 500 leAlaMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaIleAlaIleAlaI 520  
Db 29382 TGACCGCGGGGCGATGCTGTTGGCGGAGCGCTGGCATGCCCTTGTTCAGCGCGGCGT 29441  
Qy 520 al 520  
Db 29442 TA 29443  
RESULT 4  
US-09-567-969-1  
; Sequence 1, Application US/09567969  
; Patent No. 6355457  
; GENERAL INFORMATION:  
; APPLICANT: Schupp, Thomas  
; APPLICANT: Ligon, James  
; APPLICANT: Molnar, Istvan  
; APPLICANT: Zirkle, Ross  
; APPLICANT: Cyr, Devon  
; APPLICANT: Geerlach, Joern  
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES  
; FILE REFERENCE: 4-30582A  
; CURRENT APPLICATION NUMBER: US/09/567,969  
; CURRENT FILING DATE: 2000-05-10  
; PRIOR FILING DATE: 1999-06-17  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 68750  
; TYPE: DNA  
; ORGANISM: Sorangium cellulosum  
US-09-567-969-1  
Alignment Scores:  
Pred. No.: 0.047 Length: 68750  
Score: 152.50 Matches: 106  
Percent Similarity: 29.68% Conservative: 43  
Best Local Similarity: 21.12% Mismatches: 164  
Query Match: 5.24% Indels: 190  
DB: 4 Gaps: 14  
US-10-030-884-14 (1-573) x US-09-567-969-1 (1-68750)

Qy 158 TyrArgAlaAlaAlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSerIle 177  
Db 28267 ---CGGGCGGCTGATGACGGGCTCTCGCGCGCGCGGATGGTCTCGCTCGGAGC--- 28320  
Qy 178 ValSerPheArgValAspSerAspValValSerLeuAlaArgGlyAspValGluLeuGlu 197  
Db 28320 ----- 28320  
Qy 198 AlaGluProAspGlyValAlaAlaGlyAlaGlyAlaVal----- 209  
Db 28321 GCCGAGCGCGAGGTGGCGCGCGCGCTGGCGCGCGCGCGCGCGCTCGGTGCTGATCGCGG 28380  
Qy 210 -----SerSerArgGlyGlyAspAlaGlyArg-----ValArg 220  
Db 28381 GGTCAATGGCGCGAGCAGGTGCTGATCGCGCGCTGGAGCAAGCGGTGCGAGCGATCGC 28440  
Qy 221 ValThrValArgLysSerThrSerSerArgSerGluAlaAlaCys-SerHis---SerHi 239  
Db 28441 GCGGGGTTTCG 28500  
Qy 239 sSerGlnThrMetGlnPro-----ArgValSerAsnLeuSerGlyVa 253  
Db 28501 CTCGCCGCTGATGGAAACCGATGCTGGAGGAGTTCGGCGCGGTGGCGGC-GTCGGTGACGT 28559  
Qy 253 lGluIleTyrSerLeuGlnSerSerArgAsnProThrProArgGlySerSerPheAsnHi 273  
Db 28560 ACCGCGCGCCAGCGCTTTCGCTGGTGGACCACTGAGCGGGAAGGTGCTGCG---GGACG 28616  
Qy 273 sAlaAspPhePheAsnIleValGlyAlaAlaLysGlyGlyGlyAlaAlaGlyAs 293  
Db 28617 AGCTGAGCGCGCGCGGCTACTGGTGGCGCACGTGGCGGAGCGGTGCGCTTCGCGGACG 28676  
Qy 293 pGluGluLysGlyAlaCysGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 313  
Db 28677 GGGTGAAGCGCTGCACGAAGCGCGGTGCGCGGCGACGTTCTCGAAGTGGGCGCGGACCGCA 28736  
Qy 313 lAlaValProAlaLysArgLysAspLeuHisMetLeuValTrpSerSerAlaSerPr 333  
Db 28737 CGTGTCT----- 28743  
Qy 333 oValSerGluArgAlaAlaValHisValPheGlyAlaGlyAlaAspHisAlaAspVa 353  
Db 28744 -----CGGGCTGTTGCCAGCCTGCTGCGGAGCGGAGCGGAGCTGCTGCGCGT 28793  
Qy 353 lLeuAlaLysGlyAlaGlnAlaTyrAspGluTyrGlyArgAspAspTyrSerSerArgTh 373  
Db 28794 CGTTCGCGC-----CGGCG 28808  
Qy 373 rLysAsnGlySerGlyAlaAlaAspLysGlyGlyProThrLeuSerLysLeuGlySerAs 393  
Db 28809 GCGAGGAGGCTGCGGGGTGCTCGAGCGCTGGGCGAGGCTGTGGCGCGCGCGCGCTCGG 28868  
Qy 393 nSerThrAlaGlnLeuTyrProLysAspAspGlyGlyGlyArgAlaAlaValAlaMe 413  
Db 28869 TCAGTGGCGCGGCTGCTTCCCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 28928  
Qy 413 tProProAlaSerValMetThr----- 420  
Db 28929 GGCAGCGCGAGCGGTACTGGCCCGCGACATCGAGCGCTGACAGCGCTGCGCCAGCGCGCGG 28988  
Qy 421 -----ArgLeuIleLeuIle-MetValTrpArgLysLeuIleArgAsn----- 434  
Db 28989 ATCCGACCCAAAGCTGTTCTATCGCGTGAGCTGCGCGGAGATACCTCGCAGCCTCCAGA 29048  
Qy 434 ----- 434  
Db 29049 AATCAGAGGAGCGAGCGCGGAGCTGGTGGTATTGGCGGATAAGGTGGAGTCGGCG 29108  
Qy 435 -----ProA 436  
Db 29109 AGCGGTGCTGCGAGCGCTGTCGACACGTGGACTTCCATGCTGCTGCTCCATGCGCGG 29168  
Qy 436 snThrTyrSerSerLeuIleGlyValVal-----Trps 447

```

Db 29169 CAGAGACATCCGCGACCGCGAGCTGGTACCGAGGCTCGCGGCTCGAAGCGATTGGC 29228
QY 447 erLeuValSerTyrArgTgLyLeGluMetProAlaIleAlaAraSerIleSerI 467
Db 29229 AGGTAGTGTCTACTGTGGGTCTGACCGCTGTCGAGCCCTGTCGCGGCGCTCATCATG 29288
QY 467 leLeuSerAspAlaGly-----LeuGlyMetAlaMetPheSerL 480
Db 29289 AGATCGCGGACGACCGCTGCTACCGCGCGGCTGCTCGGCTTGGCTGCTGCTG 29347
QY 480 euGlyLeuPheMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAlaI 500
Db 29348 -----AGCACCGGTGCTTGTGCGCCGACTCTGGTCTG 29381
QY 500 leAlaMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaAlaSerIleAlav 520
Db 29382 TGACCCGGGGGATGATGCTTGGCGACGAGCTCGGATCGCCCTTGTGACGGCGGT 29441
QY 520 al 520
Db 29442 TA 29443
RESULT 5
US-09-568-480-1
; Sequence 1, Application US/09568480
; Patent No. 6355458
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goslach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/568,480
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 68750
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
US-09-568-480-1
Alignment Scores:
Pred. No.: 0.047 Length: 68750
Score: 152.50 Matches: 106
Percent Similarity: 29.68% Conservative: 43
Best Local Similarity: 21.12% Mismatches: 164
Query Match: 5.24% Indels: 190
DB: 4 Gaps: 14
US-10-030-884-14 (1-573) x US-09-568-480-1 (1-68750)
QY 118 ValMetGlyIleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeuMet 137
Db 28213 GTGGTGGCGGGGTGTCTCGCTGGAGATGGGTGAGGCTCGTGGCGGCGG----- 28266
QY 138 ValGlnValValLeuGlnCysIleIleTyrThrLeuMetLeuPheLeuGlu 157
Db 28266 ----- 28266
QY 158 TyrArgAlaAlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSerIle 177
Db 28267 ---CGGGCGGTGATGACAGGGGCTCTCGGGGGGGCGCGATGCTGCTCGAGC--- 28320
QY 178 ValSerPheArgValAspSerAspValValSerLeuAlaArgGlyAspValGluLeuGlu 197
Db 28320 ----- 28320

```

```

QY 198 AlaGluProAspGlyValAlaGlyAlaGlyAlaVal----- 209
Db 28321 GCGGAGGCGGAGGTGCGGCGCGGTGGCGCGCACCGCGCTCGGTGCGATCGCGCG 28380
QY 210 -----SerSerArgGlyGlyAspAlaGlyArg-----ValArg 220
Db 28381 GGTCAATGGCGCGGAGCAGGTGTGATCGCGGCGTGGAGCAAGCGGTGCGAGCGCATCGC 28440
QY 221 ValThrValArgLysSerThrSerSerArgSerGluAlaAlaCys-SerHis---SerHis 239
Db 28441 GCGGGGTTCGGCGCGCGCGCGCGCGCCGCGCACCAAGCGGTGATGCTCTCGCACCGCTCCCA 28500
QY 239 sSerGlnThrMetGlnPro-----ArgValSerAsnLeuSerGlyVa 253
Db 28501 CTCGCGCTGATGGAACCGATGCTCGAGAGTTCGGCGGGTGGCGGC-GTCGGTCACT 28559
QY 253 lGluIleTyrSerLeuGlnSerSerArgAsnProThrProArgGlySerSerPheAsnHi 273
Db 28560 ACCGGCGGCAAGCGTTTCGTGTCAGCAACCTGAGCGGGAAGTGTGCG---GGAGC 28616
QY 273 sAlaAspPhePheAsnIleValGlyAlaAlaAlaLysGlyGlyGlyAlaAlaGlyAs 293
Db 28617 AGCTGAGCGCGCGGGGTACTGGGTGCGGCACGTGCGGAGCGGTGCGCTTCGGCGACG 28676
QY 293 pGluGlnLysGlyAlaCysGlyGlyGlyGlyGlyHisSerProGlnProGlnAlaVa 313
Db 28677 GGGTGAAGCGCTGCACGAAGCGGTGCGGCACCTTCGTCAAGTGGCGCCGAGCCGA 28736
QY 313 lAlaValProAlaLysArgLysAspLeuHisMetLeuValTrpSerSerAlaSerPr 333
Db 28737 CGCTGCT----- 28743
QY 333 oValSerGluArgAlaAlaValHisValPheGlyAlaGlyAlaAspHisAlaAspVa 353
Db 28744 -----CGGGCTGTTCAGCTCCCTCCCGAGCGGAGCGGACGCTGCTGGCGT 28793
QY 353 lLeuAlaLysGlyAlaGlnAlaTyrAspGluTyrGlyArgAspAspTyrSerSerArgTh 373
Db 28794 CGTTCGCGCG-----CGGCG 28808
QY 373 rLysAsnGlySerGlyAlaAspLysGlyGlyProThrLeuSerLysLeuGlySerAs 393
Db 28809 GCGAGAGGCTGCGGGGTGCTCGAGGCGCTGGGCGGCTGGGCGCGCGCGCTCGG 28868
QY 393 nSerThrAlaGlnLeuTyrProLysAspAspGlyGlyArgAlaAlaValAlaMe 413
Db 28869 TCAGCTGCGCGCGGTCTTCCACGCGTGGCGGCGGTGCGCTGCGACCTATCCGT 28928
QY 413 tProAlaSerValMetThr----- 420
Db 28929 GGCAGCGGACGCGGTACTGGCGCGACATCGAGCCTGCGCCACGCGCGCGCG 28998
QY 421 -----ArgLeuIleLeuIle-MetValTrpArgLysLeuIleArgAsn----- 434
Db 28989 ATCCGACCAAGGCTGGTTCATCGCTGAGTGGCGGAGATACCTCGAGCTCCAGA 29048
QY 434 ----- 434
Db 29049 AATCAGAGAGCGAGCGCGGAGCTGGCTGTATTGGCGGATAAGGGTGGAGTCCGCG 29108
QY 435 -----ProA 436
Db 29109 AGCGGTGCTGAGCGGTGTCGACACGTGACTTCCATGCGTGTGTCATGCGCGCG 29168
QY 436 snThrTyrSerSerLeuIleGlyValVal-----Trps 447
Db 29169 CAGAGACATCCGCGACCGCGAGCTGGTGTGACCGAGGCTCGCGGCTCGAAGCGATTGGC 29228
QY 447 erLeuValSerTyrArgTgLyLeGluMetProAlaIleAlaAraSerIleSerI 467
Db 29229 AGGTAGTGTCTACTGTGGGTCTGGACGCGCTCGTGGTGGCGCGCTCGATCGATG 29288

```

Qy	467	leIeuSerAspAlaGly-----LeuGlyMetAlaMetPheSerL	480
		:::	:::
Db	29289	AGATCGCGACACGACCCGTCGTGTACGCGCCGTCGTGGCTTGGCTCGGTTTCGTG-	29347
Qy	480	acutGlyLeuPheMetAlaLeuGlnProArgIleAlaCysGlyAsnLysLeuAlaAlaI	500
Db	29348	-----AGCACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG	29381
Qy	500	leAlaMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaAlaSerIleAlaV	520
		:::	:::
Db	29382	TGACCCGGGGGGCATGCATCGTTGGCGACGAGCCTCGCATCGCCCTTGTTCAGCGCGCGT	29441
Qy	520	al	520
Db	29442	TA	29443
RESULT 6			
US-09-568-486-1			
; Sequence 1, Application US/09568486			
; Patent No. 6355459			
; GENERAL INFORMATION:			
; APPLICANT: Schupp, Thomas			
; APPLICANT: Ligon, James			
; APPLICANT: Molnar, Istvan			
; APPLICANT: Zirkle, Ross			
; APPLICANT: Cyr, Devon			
; APPLICANT: Goerlach, Joern			
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES			
; FILE REFERENCE: 4-30582A			
; CURRENT APPLICATION NUMBER: US/09/568,486			
; CURRENT FILING DATE: 2000-05-10			
; PRIOR APPLICATION NUMBER: 09/335,409			
; PRIOR FILING DATE: 1999-06-17			
; NUMBER OF SEQ ID NOS: 30			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 1			
; LENGTH: 68750			
; TYPE: DNA			
; ORGANISM: Sorangium cellulosum			
US-09-568-486-1			
Alignment Scores:			
Pred. No.: 0.047			
Score: 152.50			
Percent Similarity: 29.68%			
Best Local Similarity: 21.12%			
Query Match: 5.24%			
DB: 4			
US-10-030-884-14 (1-573) x US-09-568-486-1 (1-68750)			
Qy	118	ValMetGlyIleProLeuLeuArgGlyMetTyrGlyAlaSerSeraGlyThrLeuMet	137
		:::	:::
Db	28213	GTGCGTGGCGGGGTGTTCTCGCTGGAAGATGGGTGAGGCTCGTGGCGCGCG-	28266
Qy	138	ValGlnValValLeuLeuGlnCysIleIleTrpTyrThrLeuMetLeuPheLeuGlu	157
Db	28266	-----	28266
Qy	158	TyrArgAlaAlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaSerIle	177
Db	28267	---CGGCGGCTGATGACAGGGGCTCTCGCGCGCGCGGATGTTGTCGCTCGAGC---	28320
Qy	178	ValSerPheArgValAspSerAspValValSerLeuAlaArgGlyAspValGluLeuGlu	197
Db	28320	-----	28320
Qy	198	AlaGluProAspGlyValAlaGlyAlaGlyAlaVal	209
Db	28321	GCCGGAGCGGAGGTGGCGCGCGGTGGCGCGCGCGCGCGCGTGGTGTGATCGCGCG	28380
Qy	210	-----SerSerArgGlyGlyAspAlaGlyArg	220
		:::	:::

28381	DB	GETCAATGGCGCGGACGAGTGGTGTATCGCGCGGTGGAGCAACGGGTGCAGCGATCGC	28444
221	QY	ValThrValArgLysSerThr-SerSerArgSerGluAlaAlaCys-SerHis---SerHis	239
28441	DB	GCGCGGGTTCGCGCGCGCGCGCGCGCCACCAAGCGCTGCATGTCTCGACGGCTCCCA	28500
239	QY	sSerGlnThrMetGlnPro-----ArgValSerAenLeuSerGlyVa	253
28501	DB	CTCGCGCTGATGGAACCGATGCTGGAGGAGTTGCGGCGGTGGCGGC-GTCTGGTACGT	28559
253	QY	lGluIleTyrSerLeuGlnSerSerArgAsnProThrProArgGlySerSerPheAenHi	273
28560	DB	ACCGCGCCCAAGCGTTTCGTCTGGTGGACAACCTGACGCGGAAGGTGTCTCG---GGACG	28616
273	QY	salasapPhePheAenIleValGlyAlaAlaAlaLysGlyGlyGlyAlaAlaGlyAs	293
28617	DB	AGCTAGCGCGCGCGGGTACTGGGTGCGGCACCTGCGGAGGCGGTGCGCTTCGCGACG	28676
293	QY	pGluGlnLysGlyAlaCysGlyGlyGlyGlyGlyGlyGlyHisSerProGlnProGlnAlaVa	313
28677	DB	GGGTGAAGCGCTGCACGAAGCCGCTGCGGCACGTTCTCGAAGTGGCCGCGAAGCCGA	28736
313	QY	lalaValProAlaLysArgLysAspLeuHisMetLeuValTrpSerSerAlaSerPr	333
28737	DB	CGTGTCT-----	28743
333	QY	oValSerGluArgAlaAlaAlaValHisValPheGlyAlaGlyAlaAspHisAlaAspVa	353
28744	DB	-----CGGGCTGTGCCAGCTGCCTGCCGAGCGGAGCGGACGCTGTGGCGT	28793
353	QY	lLeuAlaLysGlyAlaGlnAlaTyrAspGluTyrGlyArgAspTyrSerSerArgTh	373
28794	DB	CGTTCGCGC-----CGGCG	28808
373	QY	rLysAenGlySerGlyAlaAspLysGlyProThrLeuSerLysLeuGlySerAs	393
28809	DB	GGCAGAGGCTCGGGGTGCTCGAGCGCTGGCAGCGTGGCGCGCGCGCGCTCG	28868
393	QY	nSerThrAlaGlnLeuTyrProLysAspGlyGlyGlyArgAlaAlaAlaValAlaMe	413
28869	DB	TCAGCTGGCGCGGCTCTCCCGACGGCTGGCGCGGTCGCTGCCACCTATTCGT	28928
413	QY	tProProAlaSerValMetThr-----	420
28929	DB	GGCAGCGGACGGGTACTGGCCGACATCGAGCTGCACCGCTGCCACGACGCGCG	28988
421	QY	-----ArgLeuIleLeuIle-MetValTrpArgLysLeuIleArgAen-----	434
28989	DB	ATCCGACCAAGCTGGTCTATCGCTGAGATGTCGCGGAGATCTCGCAGCTCCAGA	29048
434	QY	-----	434
29049	DB	AATCAGAGAGGCGACCGCGGAGCTGGCTGGTATTGGCGGTAAGGTGGAGTCCGCG	29108
435	QY	-----ProA	436
29109	DB	AGCGGTCTCGACGCTGTCCACACGCTGGACTTCCATGCGTCTGTCTCCATGCGCGG	29168
436	QY	snThrTyrSerSerLeuIleGlyValVal-----TrpS	447
29169	DB	CAGAGACATCCGCGACCGCGAGCTGGTGACCGAGGCTCCCGCGGTTCGAAGCGATTGC	29228
447	QY	erLeuValSerTyrArgTrpGlyIleGluMetProAlaIleIleAlaArgSerIleSerI	467
29229	DB	AGGTAGTGTCTACTCTGTTGGGTCTGGAGCCGCTCGTGGTGGAGGGCTCGATCGATG	29288
467	QY	leLeuSerAspAlaGly-----LeuGlyMetAlaMetPheSerL	480
29289	DB	AGATCGGACGCGGACCGCTCGTGTACCGCGCGGTGCTCGCTGGCTCGGTTCGTG-	29347
480	QY	eGlyLeuPheMetAlaLeuGlnProArgIleIleAlaCysGlyAenLysLeuAlaAlaI	500
29348	DB	-----AGACCGCTGTTGTCGCCCGACATCGGGTCTG	29381





Db 29442 TA 29443

RESULT 8

US-09-567-899-1

; Sequence 1, Application US/09567899

; Patent No. 6383787

; GENERAL INFORMATION:

; APPLICANT: Schupp, Thomas

; APPLICANT: Ligon, James

; APPLICANT: Molnar, Istvan

; APPLICANT: Zirkle, Ross

; APPLICANT: Cyr, Devon

; APPLICANT: Goerlach, Joern

; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES

; FILE REFERENCE: 4-30582A

; CURRENT APPLICATION NUMBER: US/09/567,899

; PRIOR FILING DATE: 2000-05-10

; PRIOR FILING DATE: 1999-06-17

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1

; LENGTH: 68750

; TYPE: DNA

; ORGANISM: Sorangium cellulosum

US-09-567-899-1

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:
152.50	106	68750	
Percent Similarity:	29.68%	Conservative:	43
Best Local Similarity:	21.12%	Mismatches:	164
Query Match:	5.24%	Indels:	190
	4	Gaps:	14

US-10-030-884-14 (1-573) x US-09-567-899-1 (1-68750)

Qy 118 ValMetGlyLeuProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeuMet 137

Db 28213 GTGCGTGGCGGGGTGTTCTCCTGGAAGATGGGTGAGCTCGTGGCGGCGG 28266

Qy 138 ValGlnValValLeuGlnCysIleTyrThrLeuMetLeuPheLeuGlu 157

Db 28266 28266

Qy 158 TyrArgAlaAlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaSerIle 177

Db 28267 ---CGGCGGCTCATGACAGGGCTCTCGCGGGCGCGGCGGATGGTCTCGGAGC 28320

Qy 178 ValSerPheArgValAspSerAspValValSerLeuAlaArgGlyAspValGluLeuGlu 197

Db 28320 28320

Qy 198 AlaGluProAspGlyValAlaGlyAlaVal 209

Db 28321 GCCGAGGCGGAGGTGGCGGCGGCTCGCGCGCACGCGCGTGGTGTGATCGCGCGC 28380

Qy 210 ---SerSerArgGlyGlyAspAlaGlyArg 220

Db 28381 GGTCAATGGCGGACGACAGTGTGTATCGCGGGCTGGAGCAAGCGGTCCAGCGCATCGC 28440

Qy 221 ValThrValArgLysSerThrSerSerArgSerGluAlaAlaCys-SerHis 239

Db 28441 GCGCGGGTTCGCGCGCGCGCGCGCGCCACCAAGCGGCTGCATGTCTCGCACGCGTCCCA 28500

Qy 239 sSerGlnThrMetGlnPro 253

Db 28501 CTCGCGCTGATGGACCGATGTGTGAGGAGTTCGGCGGGGTGGCGGC-GTCGGTGCAGT 28559

Qy 253 lGluIleTyrSerLeuGlnSerSerArgAsnProThrProArgLysSerPheAsnHi 273

Db 28560 ACCGGCGCCCAAGCGTTTCCTGCTGTGAGCAACCTGAGCGGGAAGTGTGTCG 28616

Qy 273 sAlaAspPhePheAsnIleValGlyAlaAlaAlaLysGlyGlyGlyAlaAlaGlyAs 293

Db 28617 AGCTGAGCGCGCGGGTACTGGTGGCGCACGTGGGAGCGGTGCGCTTCGCGAGC 28676

Qy 293 pGluGluLysGlyAlaCysGlyGlyGlyGlyGlyHisSerProGlnProGlnAlaVa 313

Db 28677 GGGTGAAGCGCGTGCACGAAGCGCGGTGGCGCACGTTCGTAAGTGGCGCCGAGCCGA 28736

Qy 313 lAlaValProAlaLysArgLysAspLeuHisMetLeuValTrpSerSerAlaSerPr 333

Db 28737 CGCTGCT 28743

Qy 333 oValSerGluArgAlaAlaValHisValPheGlyAlaGlyAlaAspHisAlaAspVa 353

Db 28744 ---CGGCTGTTGCCAGCTGCTCGCGAGCGCGAGCGAGCGCTGCTGGCGT 28793

Qy 353 lLeuAlaLysGlyAlaGlnAlaTyrAspGluTyrGlyArgAspAspTyrSerSerArgTh 373

Db 28794 CGTTGGCGC 28808

Qy 373 rLysAsnGlySerGlyAlaAspLysGlyProThrLeuSerLysLeuGlySerAs 393

Db 28809 GCGAGGAGGCTGGCGGGTGTCTGAGCGCTGGGCGAGGTGGCGCGCGCGGCGTGG 28868

Qy 393 nSerThrAlaGlnLeuTyrProLysAspAspGlyGluGlyArgAlaAlaAlaValAlaMe 413

Db 28869 TCAGCTGGCGCGGCTCTTCCACCGCTGGCGGGGTGGCGCTGCCGCTACCTACCT 28928

Qy 413 tProProAlaSerValMetThr 420

Db 28929 GGCAGCGCGCAGCTGCTGCGCGACATCGAGCTGACAGCGCTGCCAGCGCGCGCGG 28988

Qy 421 ---ArgLeuIleLeuIle-MetValTrpArgLysLeuIleArgAsn 434

Db 28989 ATCCGACCCAGGCTGTCTATCGCGTGGACTGGCGCGGAGATACCTCGCAGCTCCAGA 29048

Qy 434 434

Db 29049 AATCAGAGGAGCGAGCGCGGAGTGGTGTATTGGCGGATAGGTGGAGTGGCGG 29108

Qy 435 435

Db 29109 AGCGGTGCTGCGAGCGGTGTCACAGTGGACTTCATGCGTGGTCCATGCGCGCGG 29168

Qy 436 snThrTyrSerSerLeuIleGlyValVal 447

Db 29169 CAGAGACATCGCGACGCGCGAGCTGTGACCGAGGTGCGCGCGGTTCGAAGCGATTGGC 29228

Qy 447 erLeuValSerTyrArgTrpGlyIleGluMetProAlaIleIleAlaArgSerIleSerI 467

Db 29229 AGTAGTGTCTACCTGTGGGTCTGAGCGCTGCGAGCGCTGCGTGGTGGGAGGCGTTCATG 29288

Qy 467 leLeuSerAspAlaGly 480

Db 29289 AGATCGCGAGCGACCCCGTGTACCGCGCGGTGCTCGCTGGCTGGCTTCGTTCTG 29347

Qy 480 euGlyLeuPheMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAlaAla 500

Db 29348 ---AGCACCGTGTCTTGTTCGCCCGCGACTCTGGGTGCG 29381

Qy 500 leAlaMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaAlaSerIleAlaV 520

Db 29382 TGACCGCGGGGCGCATGTCATGTTGGCGAGCGCTGCGATCGCCCTTGTGTGCGCGCG 29441

Qy 520 al 520

Db 29442 TA 29443

RESULT 9

US-09-443-501A-2

; Sequence 2, Application US/09443501A

; Patent No. 6303342

; GENERAL INFORMATION:





; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; FILE REFERENCE: TUBERCULOSIS  
; CURRENT APPLICATION NUMBER: US/09/103,840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1: 4411529  
; LENGTH: 4411529  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; OTHER INFORMATION: H37Rv  
US-09-103-840A-1

Alignment Scores:  
Score: 588 Length: 4411529  
Pct Ident: 139.00 Matches: 111  
Best Local Similarity: 31.87% Conservative: 49  
Query Match: 22.11% Mismatches: 191  
DB: 4.78% Indels: 152  
Gaps: 21

US-10-030-884-14 (1-573) x US-09-103-840A-1 (1-4411529)

QY 77 LeuGlnLysValAlaValLeuAlaLeuLeuAlaLeuAlaSerArgGlyLeuSerSerPro 96  
DB 363544 TTGGACACAGCGGACGATAGAGCTCGTTCACATGTCGGCGCGCGCGATGTCGCG 363485  
QY 97 ArgAlaLeuGlyLeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThr 116  
DB 363484 GGAGCGATCAAC-----TTCCGCGCGCGCGGTTCGCT----- 363452  
QY 117 LeuValMetGlyIleProLeuLeuArgGlyMetTyGlyAlaSerSerAlaGlyThrLeu 136  
DB 363451 -----GGTACACCGCTGCGT-----AGGCATCG-ATGTCCACGCGCGGTACGTT 363405  
QY 137 MetVal-----GlnValValValLeuGlnCysIleIleTrpTyThrLeuMetLeu 153  
DB 363404 TTAGCGGTGTATTCGCGAGACGACACGCGGTCTCGTTG----- 363363  
QY 154 PheLeuPheGluTyArgAlaAlaAlaArgAlaLeuValLeuAspGlnPheProAspGlyAla 173  
DB 363362 -----CGCAGTGCTGCTCAATCATTCCGACCGATTCAAAACCGCGCGC 363318  
QY 174 AlaAlaSer-----IleValSer 179  
DB 363317 CTGCGCTCGCGCTCGATCCGCGAGGACGCGCGCTCCACGTCACGAGATCTTCTCT 363258  
QY 180 PheArgValAspSerAspValValSerLeuAlaArgGlyAspValGluLeuGluAlaGlu 199  
DB 363257 TTGTCATACGAGAGATCTCTACTATCTGAGACGC----- 363222  
QY 200 ProAspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyAspAlaGlyArgVal 219  
DB 363221 -----GGTGTGCGCGCACCGCG----- 363204  
QY 220 ArgValThrValArgLysSerThrSerSerArgSer---GluAlaAlaCysSerHisSer 238  
DB 363203 -----AAACGGGATTCCCGCGGTCTGTCTATCGAGCGCTGTAGTTGGCGC 363159  
QY 239 HisSerGlnThrMetGlnProArgValSerAsnLeuSerGlyValGluIleTySerLeu 258  
DB 363158 ACAAACTGCAATGCGGACCGCGGTATCTACA----- 363126  
QY 259 GlnSerSerArgAsnProThrProArgGlySerSerPheAsnHisAlaAspPhePheAsn 278  
DB 363125 ---GACTCGGTGAATCCGACTCAGCCGG----- 363099  
QY 279 IleValGlyAlaAlaAlaLysGlyGlyAlaAlaGlyAspGluGluLysGlyAla 298  
DB 363098 ---GTTGCGGTGTTGCTTTCGCGCGGTGGACCGTC-----GAGCCCTTGGCGC 363048  
QY 299 CysGlyGlyGlyGlyGlyHisSerProGlnProGlnAlaValAlaValProAlaLys 318

DB 363047 GTCCGCGCGGTGCGCCCTTGACGCGCGCGCTACCGCGCGCGCCCTTTCCGCGCGC 362988  
QY 319 ArgLysAspLeuHisMetLeuValTrpSerSerAlaSerProValSerGluArgAla 338  
DB 362987 CGC-----ACCGTCCCGCCCAACACCGCGCT 362961  
QY 339 AlaValHisValPheGlyAlaGlyGlyAla-----AspHisAlaAspValLeuAla 355  
DB 362960 ACCGCGCGTCCCGCGCGCGCGGTACCGCATTTGAAATGCGCCACCGTTGCGCGC 362901  
QY 356 LysGlyAlaGlnAlaTyArgAspGluTyGlyArgAspAspTySerSerArgThr----- 373  
DB 362900 GCGCGCGCGGTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 362841  
QY 373 ----- 373  
DB 362840 GC 362781  
QY 374 -----LysAsnGlySerGlyAlaAspLysGly-----GlyProThrLeuSerLys 399  
DB 362780 ACCGCGGTACCGCGCGCTCGTCCGCGCAATAGGGGTATTTACGAGCGCGCATCGCGCGC 362721  
QY 390 LeuGlySerAsnSerThrAlaGlnLeuTyProLysAspAspGlyGluGlyArgAlaAla 409  
DB 362720 GTGCGCGCGCATCGC 362661  
QY 410 AlaValAla-----MetProAlaSerValMetThrArgLeuIleLeuMetVal 427  
DB 362660 ACCGTGCGCGCGTTCGCGATCCACAGCGCGCATGACACCGCGCGCGCGCGCGC 362601  
QY 428 TrpArgLysLeuIleArgAsnProAsnThrTySerSerLeu-----IleGlyValValTrp 446  
DB 362600 GGTGCGCGCGCATATCGTGGCGCTAAACCGCGCGCGCGCGCGCGCGCGCGCGCGC 362541  
QY 447 SerLeuValSerTyArgTrpGlyIleGluMetProAlaIleIleAlaArgSerIleSer 466  
DB 362540 GCCTGTGC 362481  
QY 467 IleLeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeuGlyLeuPheMetAlaLeu 486  
DB 362480 GCC-----GCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 362445  
QY 487 GlnProArgIleIleAlaCysGlyAsnLys-----LeuAlaAlaIleAlaMet 502  
DB 362444 -----GCCGCGTGTGTGACCGCATCCGCGCTGACGCGCGCGCGCGCGCGCGC 362397  
QY 503 GlyValArgPheValAlaGlyProAlaValMetAlaAlaAlaSerIleAlaValGlyLeu 522  
DB 362396 GGCACCGC 362337  
QY 523 ArgGly 524  
DB 362336 AGAGGA 362331

RESULT 12  
US-08-864-038A-1  
; Sequence 1, Application US/08864038A  
; Patent No. 6001592  
; GENERAL INFORMATION:  
; APPLICANT: Kunio NAKASHIMA et al.  
; TITLE OF INVENTION: NOVEL POLYPEPTIDE GENE CDNA, VECTOR  
; TITLE OF INVENTION: CONTAINING SAID CDNA, HOST CELLS TRANSFORMED WITH SAID  
; VECTOR, POLYPEPTIDE PRODUCED THEREBY, METHOD OF PRODUCING  
; TITLE OF INVENTION: SAID POLYPEPTIDE, DNA ENCODING SAID POLYPEPTIDE AND ANTIBODY  
; TITLE OF INVENTION: TO SAID POLYPEPTIDE  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: 812-5 Hirano  
; STREET: Isshinden  
; CITY: Tsu-city  
; STATE: Mie-prefecture  
; COUNTRY: JAPAN



RESULT 14  
US-08-864-038A-4  
; Sequence 4, Application US/08854038A  
; Patent No. 6001592  
; GENERAL INFORMATION:  
; APPLICANT: Kunio NAKASHIMA et al

Qy	455	IleGluMetProAlaIleIleAlaArgSerIleSerIleLeuSerAspAlaGlyLeuGly	474
Db	1721	GGTGGAAATG-----GGTGGAGATTTCGGA	1744
Qy	475	MetAlaMet-----PheSerLeuGlyLeuPheMetAlaLeuGlnProArgIleIle	491
		::: :::	
Db	1745	GTAGGTCCTGGTGGAGATTTCGAGGAGGATT-----	1777
Qy	492	AlaCysGlyAsnLysLeuAlaIleIleAlaMetGlyValArgPheValAlaGly-----	509
Db	1778	---GGTGGTGGATCATCAGCAGCAGCTGCTGCCGTGCTGCAGCCGCCCTGGATTGGT	1834
Qy	509	-----	509
Db	1835	GGAGGTGGACGAAGAGGTAGAGGTAGCAGCAGCTGGAGGCGATGCGCAGCTAACGAGCT	1894
Qy	510	ProAlaValMetAlaAlaAlaSerIleAlaValGlyLeuArgGlyValLeuLeuHisIle	529
Db	1895	AGTGTGTGTAGTGCACCCCGCGCTGCTGCTGCTGCGAGGATCTCGTCTGATGTT	1954
Qy	530	AlaIleValGlnAlaAlaLeuProGlnGlyIleValProPheValPheAlaLysGluTyr	549
Db	1955	GCGCTGCCGCTGCTGCACCC-----GCAGCTATGTATC	1987
Qy	550	GlyValHisProAspIleLeuSerThrAlaTyrGlyProIleThrSerHisGlyPhe	568
Db	1988	GGTGACGGTGCTGAT-----GGACCTGATTTCGATAATGGAATTC	2026

RESULT 14  
 US-08-864-038A-4  
 ; Sequence 4, Application US/08864038A  
 ; Patent No. 6001592  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kunio NAKASHIMA et al.







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 3, 2004; 11:11:45; Search time 1319 Seconds  
(without alignments)  
1588.862 Million cell updates/sec

Title: US-10-030-884-14

Perfect score: 2310

Sequence: 1 MITALDIYHLTVVPLVYA.....DILSTAYGPTSHGFITCHS 573

Scoring table:

BLOSUM62

Xgapop 10.0, Xgapext 0.5

Ygapop 10.0, Ygapext 0.5

Rgapop 6.0, Rgapext 7.0

Delop 6.0, Delext 7.0

Searched: 2421054 seqs, 1828716029 residues

Total number of hits satisfying chosen parameters: 4842108

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_epool/US10030884/runat\_24022004\_103926\_2050/app\_query.fasta\_1.711  
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US10030884@cgn2\_1.1.57 @runat\_24022004\_103926\_2050  
-NCPU=6 -ICPU=3 -NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
-----					

1	2129.5	73.2	1926	15	US-10-260-238-565	Sequence 565, App
2	1822	62.6	2346	12	US-10-424-599-134028	Sequence 134028,
3	1643	56.5	3020	12	US-10-424-599-45725	Sequence 45725, A
4	1621.5	55.7	3374	12	US-10-424-599-86575	Sequence 86575, A
5	1603.5	55.1	1759	15	US-10-260-238-30	Sequence 30, Appl
6	1591.5	54.7	2415	12	US-10-425-114-16771	Sequence 16771, A
7	1560.5	53.6	1923	9	US-09-938-842A-847	Sequence 847, App
8	1542.5	53.0	2272	11	US-09-938-842A-847	Sequence 847, App
9	1542.5	53.0	2272	12	US-10-425-114-1349	Sequence 1349, Ap
10	1542	53.0	2276	12	US-10-425-114-1349	Sequence 1349, Ap
11	1536	52.4	1860	9	US-09-938-842A-848	Sequence 848, App
12	1526	52.4	1860	11	US-09-938-842A-1305	Sequence 1305, Ap
13	1406.5	48.3	1983	9	US-09-938-842A-1305	Sequence 1305, Ap
14	1406.5	48.3	1983	11	US-09-938-842A-1305	Sequence 1305, Ap
15	1272.5	43.7	1928	12	US-10-425-114-10077	Sequence 10077, A
16	1147	39.4	1744	12	US-10-425-114-7737	Sequence 7737, Ap
17	1027	35.3	1273	12	US-10-425-114-15739	Sequence 15739, A
18	1004.5	34.5	1334	12	US-10-425-114-2319	Sequence 2319, Ap
19	1004	34.5	1321	12	US-10-425-114-33502	Sequence 33502, A
20	1003	34.5	1199	12	US-10-425-114-29155	Sequence 29155, A
21	1002.5	34.5	1313	12	US-10-425-114-11838	Sequence 11838, A
22	1000	34.4	1091	12	US-10-425-114-3558	Sequence 3558, Ap
23	987	33.9	651	15	US-10-260-238-5337	Sequence 5337, Ap
24	946.5	32.5	1539	12	US-10-425-114-29226	Sequence 29226, A
25	937.5	32.2	1464	12	US-10-424-599-104351	Sequence 104351,
26	892.5	30.7	1146	12	US-10-424-599-59026	Sequence 59026, A
27	825	28.4	1271	12	US-10-424-599-18686	Sequence 18686, A
28	795.5	27.3	1736	12	US-10-424-599-134033	Sequence 134033,
29	772.5	26.5	1280	12	US-10-424-599-45727	Sequence 45727, A
30	767.5	26.4	1122	15	US-10-424-599-134029	Sequence 134029,
31	757	26.0	702	15	US-10-260-238-5464	Sequence 5464, Ap
32	733.5	25.2	1211	12	US-10-424-599-45726	Sequence 45726, A
33	611	21.0	971	12	US-10-424-599-62375	Sequence 62375, A
34	590	20.3	1045	15	US-10-260-238-4000	Sequence 4000, Ap
35	568.5	19.5	761	12	US-10-425-114-31117	Sequence 31117, A
36	561	19.3	987	12	US-10-425-114-13089	Sequence 13089, A
37	521	17.9	729	12	US-10-424-599-59027	Sequence 59027, A
38	513	17.6	673	12	US-10-424-599-71900	Sequence 71900, A
39	465	16.0	606	12	US-10-424-599-33427	Sequence 33427, A
40	447	15.4	558	12	US-10-424-599-52693	Sequence 52693, A
41	427.5	14.7	650	12	US-10-424-599-33472	Sequence 33472, A
42	410	14.1	869	12	US-10-424-599-84904	Sequence 84904, A
43	365	12.5	644	12	US-10-425-114-30762	Sequence 30762, A
44	328	11.3	574	15	US-10-260-238-5138	Sequence 5138, Ap
45	323	11.1	2000	15	US-10-260-238-2559	Sequence 2559, Ap

#### ALIGNMENTS

RESULT 1  
US-10-260-238-565  
; Sequence 565, Application US/10260238  
; Publication No. US20040016025A1  
; GENERAL INFORMATION:

; APPLICANT: Budworth, Paul R.  
; APPLICANT: Moughamer, Todd G.  
; APPLICANT: Briggs, Steven P.  
; APPLICANT: Cooper, Bret  
; APPLICANT: Glazebrook, Jane  
; APPLICANT: Goff, Stephen A.  
; APPLICANT: Katagiri, Fumiaki  
; APPLICANT: Krepis, Joel  
; APPLICANT: Provart, Nicholas  
; APPLICANT: Ricke, Darrell  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION  
; FILE REFERENCE: 60111-NP  
; CURRENT APPLICATION NUMBER: US/10/260,238  
; CURRENT FILING DATE: 2002-09-26  
; PRIOR APPLICATION NUMBER: US 60/325,448  
; PRIOR FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: US 60/325,277  
; PRIOR FILING DATE: 2001-09-26

```

; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 6077
; SEQ ID NO 565
; LENGTH: 1926
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: N region
; LOCATION: (261)..(261)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N region
; LOCATION: (291)..(291)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N region
; LOCATION: (486)..(486)
; OTHER INFORMATION: n = any nucleotide
US-10-260-238-565

Alignment Scores:
Pred. No.: 1,4e-211 Length: 1926
Score: 2129.50 Matches: 458
Percent Similarity: 77.58% Conservative: 23
Best Local Similarity: 73.87% Mismatches: 42
Query Match: 73.18% Indels: 97
DB: 15 Gaps: 16

US-10-030-884-14 (1-573) x US-10-260-238-565 (1-1926)
QY 1 MetIleThAlaLeuAspLeuTyrHisValLeuThrAlaValProLeuTyrValAla 20
DB 1 ATGATCACGGTGGTGACCTGTACACAGTCTTGACGGCGGTGGTGGTGTACGTGGG 60
QY 21 MetThLeuAlaTyrGlySerValArgTrrTrrArgIlePheThrProAspGlnCysSer 40
DB 61 ATGACGCTGGCGTACGCTCCGTCGGTGGTGGCGCATCTTCTCCCGGACAGTGTCTCC 120
QY 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60
DB 121 GGATCAACCGCTTCGTGGCGCTTCGCGCTTCGCGCTCTCTCTCCATTCATCTCC 180
QY 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80
DB 181 ACCAACHACCCCTTCGCCATGAACCTCCGCTCCCTCGCGCGGACACGCTCCAGAGCTC 240
QY 81 AlaValLeuAlaLeuLeuAla-----LeuAlaSerArgGlyLeuSerSerPro 96
DB 241 ATCGTCTCGCCCTCTCTCGNCTCTGGTGGCGGCTCTCCGCGCGCGG-----288
QY 97 ArgAlaLeuGlyLeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThr 116
DB 289 -----TCNCTCGACTGGCTCATCACCTCTCTCTCCCTCTCCACCTCCCAACACC 339
QY 117 LeuValMetGlyIleProLeuLeuArgGlyMetTyrGlyAlaSerSerAla-----133
DB 340 CTCGTATGGGCGATCCCGCTGTCTCAAGGGAGATGACGCGCGCGCGCGCGCGCGCGCG 399
QY 134 -----GlyThrLeuMetValGlnValValLeuGlnCysIleIleTrrTyrThr 150
DB 400 GCCGACTCCGCGACCTCATGTTGCGAGATCGTGTCTCCAGTGCATCATCTGTGTACAG 459
QY 151 LeuMetLeuPheLeuPheGluTyrArgAlaAlaArgAlaLeuValLeuAspGlnPhePro 170
DB 460 CTCATGTGTCTCTTCGAGTACCGNGCGCGCGCTTCTTGTCTATGAGCAGTTCGCG 519
QY 171 AspGlyAlaAlaAlaSerIleValSerPheArgValAspSerValValSerLeuAla 190
DB 520 GAC-----ACCGCGGCTCCATCGTCTCTTCGGGTGGATTCGACGCTCGTCTCGCTCG 576
QY 191 ArgGlyAsp-----ValGluLeuGluAlaGluProAspGlyValAlaGlyAlaGly 207

```

```

577 GCGGCGGTGGTGGAGCCGCGGAGCTGCAGGCGGAG-----612
208 AlaValSerSerArgGlyGlyAspAlaGlyArgValArgValThrValArgLysSerThr 227
613 -----GCGGAGGTGGGAGCGATGCGAGATGCGGCTCACCGTGCAGGTGCGAGTGCAG 663
228 SerSerArgSerGluAlaAlaCysSerHis-----SerHisSerGlnThrMetGln 244
664 AGCTCGCGTCCGAGCGCGGTCTCGACGGAGCGAGTGCACACTCGGAGTCCATGCGAG 723
245 ProArgValSerAsnLeuSerGlyValGluIleTyrSerLeuGlnSerSerArgAsnPro 264
724 CCGGCGGTCTCCAAACCTCTCCGCGGTGAGATTACTCGCTGCAGTCTGCGGGAACCCG 783
265 ThrProArgGlySerSerPheAsnHisAlaAspPhePheAsnIleValGlyAlaAlaAla 284
784 ACGCGCGTGGTCCAGCTTCAACCGCGGAGTCTTCAACATCTCTC-----831
285 LysGlyGlyGlyAlaAlaGlyAspGluGluGlyGlyAlaCysGlyGlyGlyGlyGly 304
832 -----GGCAACGGCAAGCAGCGGCGAGGAGAGAGGCGCC-----GCCGCGTGGCGG 879
305 GlyHisSerProGlnProGlnAlaValAlaValProAlaLysArgLysAspLeuHisMet 324
880 GGCACACTCGCGCGAGCGG-----GTGCTGGGGAAGAGGAGGACCTGCACATG 927
325 LeuValTrpSerSerAlaSerProValSerGluArgAla-----338
928 TTCGTGTGGAGCTCAAGCGCTCGCGGTGTGGAGGCGCGCGCGCGCGCAGCAGCTGGC 987
339 AlaValHisValPheGlyAlaGlyGlyAlaAspHisAlaAspValLeuAlaLysGlyAla 358
988 GCGGTGCACGTCTTTCGCGGTGGTGGCGCGGACCGCGGAC-----GCCAAGAGAGCT 1041
359 GluAlaTyrAspGluTyrGlyArgAspAspTyrSerSerArgThrLysAsnGlySerGly 378
1042 CAGCCCTATGATGAG-----TACAGTTTCGGGACACAGAAC-----1077
379 GlyAlaAspLysGlyGlyProThrLeuSerLysLeuGlySerAsnSerThrAlaGlnLeu 398
1078 -----GAGAAGACGGCGCGCGCTGTCCAGCTGGGCTCCAACTGCAGCGGCGAGCTC 1131
399 TyrProLysAspAspGlyGlyArgAlaAlaAlaValAlaMetProProAlaSerVal 418
1132 CGGCCAAGACGACGCGGAGGAGTGGCGGCG-----GCGATGCGCGCGCGGAGGGTG 1185
419 MetThrArgLeuIleLeuMetValTrpArgLysLeuIleArgAsnProAsnThrTyr 438
1186 ATGACGAGGTCTCTCTGATCATGTTGGAGGAAGCTGATCAGGAACCCCAACACTTAC 1245
439 SerSerLeuIleGlyValValTrpSerLeuValSerTyrArgTrpGlyIleGluMetPro 458
1246 TCCAGCTCTCTGTGTGTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1305
459 AlaIleAlaArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPhe 478
1306 GCCATATCCCGGTGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 1365
479 SerLeuGlyLeuPheMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAla 498
1366 AGCTAGGATGTTCATGGCATTCAGCCAGGATCATTCCTCTGTGGGAACTCCCTTGCT 1425
499 AlaIleAlaMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaAlaSerIle 518
1426 TCGTATGCCATGGCGGTTCAGGTTCCTCGTGGTCCGCGCTCATGCTGCGCGCTCATC 1485
519 AlaValGlyLeuArgGlyValLeuLeuHisIleAlaIleVal-----532
1486 GCCGTGCGACTTCGCGGGTGTCTTGCACTTGCCTTTCAGGTAAAGTGAAGTGAAGTTC 1545
532 -----532
1546 AGGCTTCTGCACCAAAATGATCTTAACTGCAAAAGAAATCTATCTGTAACCAATTCG 1605

```

QY 533 -----GlnAlaLeuProGlnGlyIleValProPheValPheAlaLysGlu 548  
 Db 1606 TCATTGAATCTCAGCGCGCTCTCTCAAGAACTCGGCCCTTGTTGTTGCCAAGGAG 1665  
 QY 549 TyrGlyValHisProAspIleLeuSerThrAlaTyrGlyProIleThrSerHisGlyPhe 568  
 Db 1666 TACAATGTTTCACTCAACATCTGAGCACAGCGTATGATCTCTTTTACCATCTTTTCTTTT 1725

## RESULT 2

US-10-424-599-134028  
 ; Sequence 134028, Application US/10424599  
 ; Publication No. US20040031072A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa Thomas J  
 ; APPLICANT: Kovalic David K  
 ; APPLICANT: Zhou Yihua  
 ; APPLICANT: Cao Yongwei  
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53223)B  
 ; CURRENT APPLICATION NUMBER: US/10/424,599  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 285684  
 ; SEQ ID NO 134028  
 ; LENGTH: 2346  
 ; TYPE: DNA  
 ; ORGANISM: Glycine max  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_92032C.1  
 US-10-424-599-134028

## Alignment Scores:

Pred. No.: 1,98e-179 Length: 2346  
 Score: 1822.00 Matches: 387  
 Percent Similarity: 70.83% Conservative: 55  
 Best Local Similarity: 62.02% Mismatches: 86  
 Query Match: 62.61% Indels: 96  
 Ds: 12 Gaps: 16

US-10-030-884-14 (1-573) x US-10-424-599-134028 (1-2346)

QY 1 MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValProLeuTyrValAla 20  
 Db 148 ATGATAACGTTTGGACTTGTACACCTTCTGACAGCGGTGGTCCGCTCTACGTGGCG 207  
 QY 21 MetThrLeuAlaTyrGlySerValArgTrpTrpArgIlePheThrProAspGlnCysSer 40  
 Db 208 ATGATCCTCGCTACGGCTCCGTGAAGTGGTGAAGATCTTCCACCCCGGACCAATGCTCG 267  
 QY 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
 Db 268 GGAATAAACCGTTTCGTGGCACTCTTCGACGTGCCCTCTCTCTCTCTCTCTCTCTCTCC 327  
 QY 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80  
 Db 328 ACCAACACCCCTACGCCATGAACATACAGTTTCATCGCGCGGACACACTCCAAAAGGCC 387  
 QY 81 AlaValLeuAlaLeuAlaLeuAlaSerArgGlyLeuSerSerProArgAlaLeuGly 100  
 Db 388 ATAGTCTCGAGTGTCTTGT 438  
 QY 101 LeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThrLeuValMetGly 120  
 Db 439 CTGTAATGTTCCATACACT 498  
 QY 121 IleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeuMetValGlnVal 140  
 Db 499 ATCCCGTTGCTGAAGGCGATGACGGTGATGACTCG--GGAACACTAATGTTGATGATTT 555  
 QY 141 ValValLeuGlnCysIleIleTrpTyrThrLeuMetLeuPheLeuPheGluTyrArgAla 160  
 Db 556 GTGGTCTTCAGTGTATCATCTGGTATACCTTATGTTGTTCTTTCTTTTCAGTATAGGGGT 615

QY 161 AlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSerIleValSerPhe 180  
 Db 616 GCGAGGCTTTTAATAGTGGAGCAGTTTCTGTAT--ACAGCGGTTTCCATTATCTCTCTTC 672  
 QY 181 ArgValAspSerAspValValSerLeu--AlaArgGlyAspValGluLeuLeuGluAlaGlu 199  
 Db 673 AAGGTTGATTCGATATATATTTTCGTTGATGGGAAGAGCCCTTCAGACGAGGCGGAG 732  
 QY 200 ProAspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGlyAspAlaGlyArgVal 219  
 Db 733 GTT-----GGTGATACGCGGAAGCTT 753  
 QY 220 ArgValThrValArgLysSerThrSerSerArgSerGluAlaAlaCysSerHisSerHis 239  
 Db 754 CATGTACCGTGAGGAAGTCGCGGAGTTCGCGTTCGAGATTTTCTTAGGCGCTCTCAT 813  
 QY 240 -----SerGlnThrMetGlnProArgValSerAsnLeuSerGlyValGluIleTyr 256  
 Db 814 GGCCCCAACCTCGGTGCTGTGACTTCGAGGCTTCGAATTTGACCAATCGGAGATTTAC 873  
 QY 257 SerLeuGlnSerSerArgAsnProThrProArgGlySerSerPheAsnHisAlaAspPhe 276  
 Db 874 TCGCTGACAGCTCGAGGAACCGACTCCGAGAGGCTCGAGTTTAAACCACCGGATTTT 933  
 QY 277 PheAsnIleVal----- 280  
 Db 934 TACTCCATGTTGATGTTAAGTAACAACAACAACAACAACAACCTGAGTATGAGTCCG 993  
 QY 281 -----GlyAlaAlaAlaLysGly 286  
 Db 994 AGGCAGAGTAATTTTGGAGGCTTTGATGAAGAAAGTGTGGGGGTGTGAGGCGTAATGTT 1053  
 QY 287 GlyGlyGlyAlaAlaGly----- 292  
 Db 1054 GGTGCTGGTGTACGTGGGTACCTCGGCTCGAATCGCGGATTTTCTCGGTGGCG 1113  
 QY 293 AspGluGluLysGlyAlaCysGlyGlyGlyGlyGlyGlyHisSerProGlnProGlnAla 312  
 Db 1114 GCGAAGAAGAAGGGTGTGAGAGTGTGTGTGAGGG----- 1149  
 QY 313 ValAlaValProAlaLysArgLysAspLeuHisMetLeuValTrpSerSerAlaSer 332  
 Db 1150 -----AAGGATCTCCACATGTTTGTGTGGAGTTCGAGTGTCTTCG 1188  
 QY 333 ProValSerGluArgAlaAlaValHisValPheGlyAlaGlyGlyAlaAspHis---Ala 351  
 Db 1189 CCGGTGCTCGAA--GGTGAATCCATGCTTCAGAGGTGGTGGTGGGATTTATGGAGT 1245  
 QY 352 AspValLeuAlaLysGlyAlaGlnAla-----TyrAspGluTyrGlyArgAsp 367  
 Db 1246 GACCAGCTTCCTGTTGGTGGGTGGCTCACCAAGAAATATGATGAGTTGTGTACGAT 1305  
 QY 368 AspTyrSerSerArgThrLysAsnGlySerGlyGlyAlaAspLysGlyGlyProThrLeu 387  
 Db 1306 GAGTTTAGCTTCGGGACAGAACCGTGTCTAATGGGTTGACAAAGGAGGCGGAGTCTT 1365  
 QY 388 SerLysLeuGlySerAsnSerThrAlaGlnLeuTyrProLysAspAspGlyGluGlyArg 407  
 Db 1366 TCAAAGCTTGGCTCGAGTTCACAGCTGAGCTTCGCCCTAAAGCTCAAGGTGAA----- 1419  
 QY 408 AlaAlaAlaValAlaMetProProAlaSerValMetThrArgLeuIleLeuIleMetVal 427  
 Db 1420 GCCAAACCTACTTCCATGCCCAACACAGTGTATGACAGGCTCATTTTGTATATGTTT 1479  
 QY 428 TrpArgLysLeuIleArgAsnProAsnThrTyrSerSerLeuIleGlyValValTrpSer 447  
 Db 1480 TGGAGGAAGCTGATTAGGAACCCCAACACATATTCAGGCTCTTTGCTCTCACTTGGTCT 1539  
 QY 448 LeuValSerTyrArgTrpGlyIleGluMetProAlaIleIleAlaArgSerIleSerIle 467  
 Db 1540 TTGATCTCAATCAATGAATGTTGTATGCCAGCAATGTTGCTAAATCGATATCAATT 1599

QY 468 LeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeuGlyLeuPheMetAlaLeuGln 487  
 Db 1600 TTATCTGATGCTGGTCTTGGGATGGCAATGTTTAGCCTTGGGTATTATTCATGGCAATGGAG 1659  
 QY 488 ProArgIleAlaCysGlyAsnLysLeuAlaAlaIleAlaMetGlyValArgPheVal 507  
 Db 1660 CCAAGATTATTCATGTCGAACTCGGTGCTTCTTGTCTATGTCAGTTCGTTTCTT 1719  
 QY 508 AlaGlyProAlaValMetAlaAlaAlaSerIleAlaValGlyLeuArgGlyValLeuLeu 527  
 Db 1720 ACTGGTCTTCAGTTCATGCTGTTGCTTCAATCGTTAGGGCTCAGGGAGTTCATATTG 1779  
 QY 528 HisIleAlaIleValGlnAlaAlaLeuProGlnGlyIleValProPheValPheAlaLys 547  
 Db 1780 CACATTGCTATTGTACAGGCTGCTGCTCCTCAGGGATTGCTCTTGTGTTGCTTAAG 1839  
 QY 548 GluTyrGlyValHisProAspIleLeuSerThrAla-----TyrGly----- 561  
 Db 1840 GAATACAGGTTTCATCCTGACATACCTAGACACCGGGGTATATATTTGGGATGCTATTGCT 1899  
 QY 562 ---ProIleThr 564  
 Db 1900 CTTCTATTACG 1911

RESULT 3  
 US-10-424-599-45725  
 ; Sequence 45725, Application US/10424599  
 ; Publication No. US20040031072A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa Thomas J  
 ; APPLICANT: Kovalic David K  
 ; APPLICANT: Zhou Yihua  
 ; APPLICANT: Cao Yongwei  
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53223)B  
 ; CURRENT APPLICATION NUMBER: US/10/424,599  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 285684  
 ; SEQ ID NO 45725  
 ; LENGTH: 3020  
 ; TYPE: DNA  
 ; ORGANISM: Glycine max  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_141290C.1  
 US-10-424-599-45725

Alignment Scores:  
 Pred. No.: 1,26e-160 Length: 3020  
 Score: 1643.00 Matches: 361  
 Percent Similarity: 66.51% Conservative: 58  
 Best Local Similarity: 57.30% Mismatches: 111  
 Query Match: 56.46% Indels: 100  
 DB: 12 Gaps: 17

US-10-030-884-14 (1-573) x US-10-424-599-45725 (1-3020)

QY 1 MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValProLeuTyrValAla 20  
 Db 755 ATGATCACCTTAACAGACTTCCACCATGTCATGACGTGCAATGGTGCCTCATGTGGCC 814  
 QY 21 MetThrLeuAlaTyrGlySerValArgTTPTPArgIlePheThrProAspGlnCysSer 40  
 Db 815 ATGATACTAGCTATGGCTCAGTGAAGTGGTGGAGATTTCCTCCCTGACCAATGCTCT 874  
 QY 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
 Db 875 GGCATCAACCGTTTGTGGCACTCTTTGCAAGTGCCTCTCTCCTTCCACTTCATAGCC 934  
 QY 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80  
 Db 935 TCCAAACACCCCTTATGAGATGAACCTTAGAGTTCCTAGCTGCTGACACCCCTTCAAAGATC 994

QY 81 AlavalLeuAlaLeuLeuAla-----LeuAlaSerArgGlyLeuSerSerPro 96  
 Db 995 ATAATACTAGTCTCTCTTGGAGCAACATCACCAAAAGGGGTGT----- 1045  
 QY 97 ArgAlaLeuGlyLeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThr 116  
 Db 1046 -----TTGGAATGGGCATAAACCCTTCTCTCTCCACCCCTCCCAACACT 1093  
 QY 117 LeuValMetClyIleProLeuLeuArgGlyMetTyrClyAlaSerSerAlaGlyThrLeu 136  
 Db 1094 TTGGTTATGGGCATCCCTTTGGCTCAAAGGATGATGGTGACTTCTCA---GGAGCCTC 1150  
 QY 137 MetValGlnValValValLeuGlnCysIleIleTyrThrLeuMetLeuPheLeuPhe 156  
 Db 1151 ATGTGCAAAATTGTGGTCTCTCCAGTGCATCATTTGGTACACCTTGATGCTCTTCTTGT 1210  
 QY 157 GluTyrArgAlaAlaAlaAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAsp 176  
 Db 1211 GAGTTAGAGGTGCCAAGATGCTCATCTCTGACGATGTCCTGAC---ACTGCTGCTCTC 1267  
 QY 177 IleValSerPheArgValAspSerAspValValSerLeu---AlaArgGlyAspValGlu 195  
 Db 1268 ATTGCTCTCCATCCATGTGGACTCTGATGTCATGTCATTTGGATGGAAGACACCATGAG 1327  
 QY 196 LeuGluAlaGluProAspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGlyAsp 215  
 Db 1328 ACTGAAGCTGAGATCAAG----- 1348  
 QY 216 AlaGlyArgValArgValThrValArgLysSerThrSerArgSerGluAlaAlaCys 235  
 Db 1349 GATGGTAACCTCCATGTCATCTGTGAGAAATCAATGCTTCAGATCAGACATCTTCTCT 1408  
 QY 236 SerHisSerHisSer---GlnThrMetGlnProArgValSerAsnLeuSerGlyValGlu 254  
 Db 1409 AGAAGGTCTCAGGGTCTCTCTCCACCACTCCACGGCCTTCCAAACCTTACCATGCTGAG 1468  
 QY 255 IleTyrSerLeuGlnSerSerArgAsnProThrProArgLysSerSerPheAsnHisAla 274  
 Db 1469 ATATACTCTTTGCAATCTCTAGAACCCCTACCCGAGAGGCTCTAGTTTCAACCACT 1528  
 QY 275 AspPhePheAsnIleValGlyAlaAlaAlaLysGlyGlyGlyGlyAlaAlaGlyAsp--- 293  
 Db 1529 GATTTCTACTCCATGATGGCTGCTGGTGCGAGAACTCAAACCTTGGTCTCTGATGTT 1588  
 QY 294 -----GluGluLysGly 297  
 Db 1589 TATGGCCTTTTCAGCTTCAAGAGGGCCAACTCCAAAGCCTTCTTAACATGATGATGATGTT 1648  
 QY 298 Ala-----CysGlyGlyGlyGlyGlyGlyHisSerProGlnProGln 311  
 Db 1649 GGGAGGCCAAAGTTTCATTACCATGCTGCTGTGGTGAACCTGGGCACTACCCCTGACCTAAC 1708  
 QY 312 -----AlaValAlaValProAlaLys 318  
 Db 1709 CTTGGCATGTTCTCTCCCTCTAATGGGTCCAAAGTGTGCTGCTAATGCTAATGCCAAG 1768  
 QY 319 Arg-----LysAspLeuHisMetLeu 325  
 Db 1769 AGGCCTAATGGCAGGCTCAGCTGAAGCCTGAGGATGGGAATAGGCACTTCATATGTTT 1828  
 QY 326 ValTrpSerSerAlaSerProValSerGluArgAlaAlaValHisValPheGlyAla 345  
 Db 1829 GTTTGGAGTTCAAAGTGTCTCCACAGTCTCTGACGCTTTGGTGGCCCATGATGATGA--- 1885  
 QY 346 GlyGlyAlaAspHisAlaAspValLeuAlaLysGlyAlaGlnAlaTyrAspGluTyrGly 365  
 Db 1886 GGAGTTCATGATCAGAAAGAGTCAAATGATGATATCCAGGAAAAGTGGAGATAAT 1945  
 QY 366 -----ArgAspAspTyrSerSerArgThrLysAsnGlySerGlyGlyAlaAsp 381  
 Db 1946 CATAGGACACTCAAGAAAGACTACCTAGAGAAAGATGAGTTTCAGCTTTGGGAATAGAGAA 2005  
 QY 382 LysGlyGlyProThrLeuSerLysLeuGlySerAsnSerThrAlaGlnLeuTyrProLys 401

```

Db      2006 ATG-----GACAGGAGAGATGATCATCAGCTTGAGGTGAG 2038
Qy      402 AspAspGlyGluGlyArgAlaAlaAlaValMetProProAlaSerValMetThrArg 421
Db      2039 AAGTTGGAGATGGAAACCAAA-----ACCATGGCTCCAGCAAGTGTGATGACAGG 2092
Qy      422 LeuIleLeuIleMetValTTPArgGlyLeuIleArgAsnProLeuThrThrYrSerSerLeu 441
Db      2093 CTTATATTGATTATGGTGTGGAGAAACTCATCAGAAACCCCAACACTACTCTAGCCTA 2152
Qy      442 IleGlyValValTTPSerLeuValSerTyrArgTTPGlyIleGluMetProAlaIleIle 461
Db      2153 ATTGGTCTCACTGGTCTCTCTGTTCATTCAAGTGGGAATGTTGAGATGCCCTGCCAATA 2212
Qy      462 AlaArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeuGly 481
Db      2213 GCAAAGTCTATCTCCATATTGTCCAGACGCGAGGGCTGGCATGGCCATGTTCAGTCTTGGT 2272
Qy      482 LeuPheMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAlaIleAla 501
Db      2273 CTCCTTCATGGTGTGGCAACCGAGGTGATAGCATGTGGAAATCCACAGCAGCTTTGGC 2332
Qy      502 MetGlyValArgPheGlyProAlaValMetAlaAlaAlaSerIleAlaValGly 521
Db      2333 ATGGCTGTGAGATTCCTTACAGGTCCAGCTGTCTATGGCAGCTGTTCCATTGCTGTGGA 2392
Qy      522 LeuArgGlyValLeuLeuHisIleAlaIleValGlnAlaAlaLeuProGlnGlyIleVal 541
Db      2393 CTCAAAGGTGTTCCTACACGTTGCCATTGTTCCAGGCAGCTCTTCCCAAGGAATTGTC 2452
Qy      542 ProPheValPheAlaLysGluTyrGlyValHisProAspIleLeuSerThrAla----- 559
Db      2453 CCATTGTCTTGTGTAAGGAATATAATGTATCATCTCTGTATTCACGACAGCTGTATT 2512
Qy      560 TyrGly-----ProIleThr 564
Db      2513 TTTCGGATGCTGATTGCTTTGCCATAACT 2542

RESULT 4
US-10-424-599-86575
; Sequence 86575, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ. ID NOS: 285684
; SEQ ID NO 86575
; LENGTH: 3374
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_49189C.1
US-10-424-599-86575

Alignment Scores:
Pred. No.:      2,56e-158      Length:      3374
Score:          1621.50      Matches:      375
Percent Similarity: 62.68%      Conservative: 60
Best Local Similarity: 54.03%      Mismatches: 94
Query Match:      55.72%      Indels:      165
DB:              12          Gaps:         21

US-10-030-884-14 (1-573) x US-10-424-599-86575 (1-3374)

Qy      1 MetIleThrAlaLeuAspLeuTyrHis-ValLeuThrAlaValProLeuTyrValAl 20
```

```

Db      335 ATGATAACCTGGGGAGATTTCTATCACGTTGATGTCGGCGGTGATTCGCTGTAGTGGC 394
Qy      20 MetThrLeuAlaTyrGlySerValArgTTPArgPheThrArgIlePheThrProAspGlnCysSe 40
Db      395 GATGATCTTAGGTACGGCTCGTGGGTGTGGAGATATTCTCGCGGACCAATGCTC 454
Qy      40 rGlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSe 60
Db      455 CGGCATAAACCCGCTTCGTCGGCATCTTCGCGGTTCGCTCTCTTTCTGTTCACTTCATCTC 514
Qy      60 rThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVa 80
Db      515 CTTGAAACATCCCTACGGCATGAACCTCGGTTCATCGTCGGGACACGCTCCAGAGAT 574
Qy      80 lAlaValLeuAlaLeuLeuAlaLeu-----AlaSerArgGlyLeuSerSerPr 96
Db      575 CATCATGCTCTTCGCGCTCGCAATTCGGACGAACCTCTCCGCGCAACGGG----- 623
Qy      96 oArgAlaLeuGlyLeuAspTTPSerIleThrLeuPheSerLeuSerThrLeuProAsnTh 116
Db      624 -----ACCTGGAGTGAATGATTACCAATTTCTCCCTCTCCATTCGCGCAACTCC--GGCTATT 730
Qy      116 rLeuValMetGlyIleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrIle 136
Db      674 GCTCGTGAATGGGAATCCCGCTGCTCATCGCCATGTACGGCAATACTCC--GGCTATT 730
Qy      136 uMetValGlnValValValLeuGlnCysIleIleTTPThrThrLeuMetLeuPheLeuPh 156
Db      731 GATGTGCAAGTGGTGGTGTGTCAGTGCATCATCTGGTACACGCTCCTCTCTCTCTCT 790
Qy      156 eGluTyrArgAlaAlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSe 176
Db      791 CGAGTACCGCGCGCGAAGCTTCTTATAATGGAACAGTTTCTCTGAA---ACGCTGCTTC 847
Qy      176 rIleValSerPheArgValAspSerAspValValSerLeu---AlaArgGlyAspValGl 195
Db      848 TATAGTCTCTTTCAAGTGTGATTCCGACGTCGTTTCGCTCGATGGAGAGGATTTTCTCGA 907
Qy      195 uLeuGluAlaGluProAspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGlyAs 215
Db      908 AACGGATGCCGAAGTC-----GGCGA 928
Qy      215 pAlaGlyArgValArgValThrValArgLysSerThrSerSerArgSerGluAlaAlaCy 235
Db      929 TGATGGGAAGCTTTCAGTGCAGCGGTAGGAAGTTCGAATGCATCTAGA----- 974
Qy      235 sSerHisSerHisSerGlnThrMetGlnProArgValSerLeuSerGlyValGluI 255
Db      975 -----AGTCTCTTATGATGACTCCGCGGCCCTTAACCTCACCGCGCGGAGAT 1024
Qy      255 eTyrSerLeuGlnSerSerArgAsnProThrProArgGlySerSerPheAsnHisAlaAs 275
Db      1025 TTACAGCTTCAGCTCGTCGCGTAAACCCAGCGCACGTGGCTCCAACCTTCAACACCGCGA 1084
Qy      275 pPheAsnIleValGlyAlaAlaAlaLysGlyGlyGlyAlaAlaGlyAsp----- 293
Db      1085 TTTTACTCCATGATGGGGGTACGCGCGCGGCAATCTAATTTTGTGCGCGGATTGTA 1144
Qy      294 -----GluGlyLysG 297
Db      1145 CTCGCTGAGTCCAGCTCACGCGGGGTACACGCGGCCCTCGAACTTCGAGGAGAACGG 1204
Qy      297 yAla----- 298
Db      1205 TCGCGCTGCTGTGCAACCAACGACGACGACGAGGCCATCAGTCCCCGAGTTTCG 1264
Qy      299 -----CysGlyGlyGlyGlyGlyHis----- 306
Db      1265 GGTTCATCCCGCGGTGCAGACAGTGGCGGGGCAATATCCGGCTCCCAATCCGGAATCT 1324
Qy      307 -----Ser-P 308
```



```
Db 1325 CTCTCTCCGGTTACGAAGAGTGTGAGCAAGAAATTCGACAGACTCAGCTCAGCTCAGC 1384
Qy 308 rGlnProGln-----AlaValAlaValProA 317
|||
Db 1385 CTCAGCCACAGATTCAAGCTCAAGTTGCTCCTCCTCCTCAACCTCAAGTAGCTCAACCTA 1444
Qy 317 lAllys-----AIGLYASPLEUHIEMETLEUVALTTPSSRS 329
|||
Db 1445 CCAACAGTGGGACAAAGCCCAACCCAGGTCTGAGAGGCTTCACATGTTTGTGGAGCT 1504
Qy 329 eRSeAlaSerProValSerGluAArgAlaAlaValHisValPhe---GlyAla-----G 346
|||
Db 1505 CCAGCAGCTGCGCGGCTCTCGAAGCGCGGCTCCAGCTGTTAGTGGGCTGATTCG 1564
Qy 346 lYGLyAlaAspHisAlaAspValLeuAlaLySGLyAlaGln----- 359
|||
Db 1565 GAGCGTCTCATCATCCGCTCGCTCGAACAAGGGGCCAAGAGATTAGGATTGGTGGT 1624
Qy 360 -----AlaTyrAspGluTyrGlyA 366
|||
Db 1625 CTGATGACCACCTCAAAATGGAGAAACCAACAAAGCTGCAGCAGAGAGAGTTTGGT 1684
Qy 366 rGAspAspTyrSerSerArgThrLySAsnGlySerGlyGly-----AlaAspLyS 383
|||
Db 1685 GTGAAGATTGAATTCAGGTAAAGAGGAGAGAACAGCAGATGAGAGGGAGAGAAAG 1744
Qy 383 lYGLyPro---ThrLeuSerLySLeuGlySerAsnSerThrAlaGlnLeuTyrProLyS 402
|||
Db 1745 CAGGACCCGGTGTCTAACAAGCTTGGATCTAGTTCACGGCGGAGCTACACCCGAAT 1804
Qy 402 sPAspGlyGluGlyArgAlaAlaAlaValAla-----MetProProlaSerValM 419
|||
Db 1805 CC-----GCCGTGCGCTTGTCTGTAAGAACACATGCTCCGCGGAGTGTA 1849
Qy 419 eThrArgLeuLeuLeuMetValTtpAqGLySLeuLeuArgAsnProAsnThrTyrS 439
|||
Db 1850 TGACTCGTCTCATCTGATTATGGTGTGAGAGAACTTATCCGCAACCCCAACACTACT 1909
Qy 439 eRSeLeuLeuGlyValValTtpSerLeuValSerTyrArgTPrpGlyLeuMetProA 459
|||
Db 1910 CTAGCCTCATTTGGTGTCTTGGTCCCTCGTGGCGTTAGGTGGCATGTGCAAAATGCCCA 1969
Qy 459 lAlleAlaAlaArgSerileSerileLeuSerAspAlaGlyMetAlaMetPheS 479
|||
Db 1970 AAATATAGAGAAATCAATCTCCATCTGCTGATCTGCTGGTCTTGGAAATGCTATGTTCA 2029
Qy 479 eRLeuGlyLeuPheMetAlaLeuGlnProArgileAlaCysGlyAsnLySLeuAlaA 499
|||
Db 2030 GCTTGGGTCTGTTATGGCTCTTCAACCTAAGATAATTGCTATGCGGAACCTGTGTGCA 2089
Qy 499 lAlleAlaMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaSerileA 519
|||
Db 2090 CATTTGCCATGCTGTTCGATTTCCTCAACAGGTCCTCGGCTCATGCGACAGCTTCCATCG 2149
Qy 519 lAValGlyLeuArgGlyValLeuLeuHisileAlaAlaileValGlnAlaAlaLeuProGln 539
|||
Db 2150 CTGTTGGCTACGTGGCACCCCTTACGGTGTAGCTATTGTTACGCTGCTCACTACCAAG 2209
Qy 539 lYleValProPheValPheAlaLySLeuTyrGlyValHisProAspIleLeuSerThra 559
|||
Db 2210 GGATTGTTCCATTGTTGTTGTTAGGAGTAAACCTCCACCCAGCCATTCTTAGTACAG 2269
Qy 559 lA-----TyrGly-----ProIleThr 564
|||
Db 2270 CGGTATATTGGGATGTTGATAGCACTACCAATTACT 2307
```

## RESULT 5

```
US-10-260-238-30
; Sequence 30, Application US/10260238
; Publication No. US20040016025A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, Paul R.
; APPLICANT: Moughamer, Todd G.
```

```
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Fumiaki
; APPLICANT: Krepis, Joel
; APPLICANT: Provart, Nicholas
; APPLICANT: Ricke, Darrell
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
; FILE REFERENCE: 60111-NP
; CURRENT APPLICATION NUMBER: US/10/260,238
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/325,448
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 6077
; SEQ ID NO 30
; LENGTH: 1759
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: N region
; LOCATION: (1584)..(1584)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N region
; LOCATION: (1587)..(1587)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N region
; LOCATION: (1602)..(1602)
; OTHER INFORMATION: n = any nucleotide
; US-10-260-238-30
```

```
Alignment Scores:
Pred. No.: 7,53e-157 Length: 1759
Score: 1603.50 Matches: 373
Percent Similarity: 68.74% Conservative: 47
Best Local Similarity: 61.05% Mismatches: 96
Query Match: 55.10% Indels: 96
DB: 15 Gaps: 21
```

```
US-10-030-884-14 (1-573) x US-10-260-238-30 (1-1759)
Qy 1 MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValValProLeuTyrValAla 20
|||
Db 1 ATGATAACGGGGCGGCGACTTCTACACGATGATGCGCGATGGTGGCGTGTACGTGGCG 60
|||
Qy 21 MetThrLeuAlaTyrGlySerValArgTTrpArgIlePheThrProAspGlnCysSer 40
|||
Db 61 ATGATCCTCGCTGACGGGTGCGTGAAGTGTGGCGCATCTTCACGCCCGCACCATGTCTCC 120
|||
Qy 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60
|||
Db 121 GGGATCAACCGCTTCGTGGGCTCTTCGGCGTCCCGCTCTCTCTCTCCATTCATCTCC 180
|||
Qy 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLySVal 80
|||
Db 181 ACCAACAAACCCCTACACCATGAACCTCCGCTTCATCGCGCGCGACACCCCTGCAGAAAGCTC 240
|||
Qy 81 AlaValLeuAlaLeuAlaLeuAlaSerArgGlyLeuSerSerProArgAlaLeuGly 100
|||
Db 241 ATGCTCTCGCTCTCTCTCCCTGTGGAGCCAC---CTCTCC-----CGCCGGGCTCC 291
|||
Qy 101 LeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThrLeuValMetGly 120
|||
Db 292 CTCGAGTGGACCATCACCTCTCTCCCTCTCCACGCTGCCAACACACGCTCGTCAATGGGG 351
|||
Qy 121 IleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeuMetValGlnVal 140
```



QY 61 ThrAsnAspPropPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80  
DB 369 ACCAAGACCCGTACGCCATGAACTCCCGTTCCTGGCCGCGACACCGCTGCAGAGCTG 428  
QY 81 AlaValLeuAlaLeuLeuAlaLeuAlaSerArg-----GlyLeuSerSerProArg 97  
DB 429 GTCGTCTCCGCGGGGTGGCCGCGTCTCGCTCCCTGCTCCAGCTGGCGCGCGCG 488  
QY 98 AlaLeuGlyLeuAspTrpSerLeuThrLeuPheSerLeuSerThrLeuProAsnThrLeu 117  
DB 489 -----CTGACTGGTTCATCAGCTCTCTCGTGTCCAGCTGCCCAACACGGTG 539  
QY 118 ValMetGlyLeuProLeuLeuArgGlyMetTrpGlyAlaSerSerAlaGlyThrLeuMet 137  
DB 540 GTGATGGCATCCCGTCTGTGATCGCATGTATCGGT-----CCCTACGCGGGTCCCTGATG 596  
QY 138 ValGlnValValValLeuGlnCysIleLeuTrpThrLeuMetLeuPheLeuPheGlu 157  
DB 597 GTGCAAGTGTGTGTCTCAGTGTATCATCTGTATCATCTGTCTGTCTCTCTCTCT 656  
QY 158 TyrArgAlaAlaAlaAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSerIle 177  
DB 657 TTCGCGCGCGCGGACGCTCATCGCGGACCATGTTCCCGACACCGCGCGGCGC---ATC 713  
QY 178 ValSerPheArgValAspSerValValSerLeuAlaArgGlyAspValGluLeuGlu 197  
DB 714 GCCTCTCCGTCACGTGACCGGACGCGGTGTCTGTGTGTGTGTGTGTGTGTGTGTGT 773  
QY 198 AlaGluProAspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGlyAlaGly 217  
DB 774 GCCGAGGTG-----ACGCGCGCGCGCTCCAACTGACGCGCGCGGAGATATCTCA 893  
QY 218 ArgValArgValThrValArgLysSerThrSerSerArgSerGluAlaAlaCysSerHis 237  
DB 795 CGGCTCCAGTGTGCGTGTGCGCGGTGTGCGCGGTGTGCGCGGTGTGCGCGGTGTGCG 851  
QY 238 SerHisSerGlnThrMetGlnProArgValSerAsnLeuSerGlyValGluIleTrpSer 257  
DB 852 -----ACGCGCGCGCGCTCCAACTGACGCGCGCGGAGATATCTCA 893  
QY 258 LeuGlnSerSerArgAsnProThrProArgGlySerPheAsnHisAlaAspPhe 277  
DB 894 TG-AGCTCGTGGCGGACGACAGCGCGGCGTCCAACTTCAACCGCGCGCTCTCTC 952  
QY 278 AsnIleVal-----GlyAlaAlaAlaLysGlyGly 288  
DB 953 GCCATGTGACG 1012  
QY 289 GlyAla-----AlaGlyAsp 293  
DB 1013 GCCGCGGAGCTGTACTCGATGCTCTGTCGCGGCGCGCGCGCGCGCGCGCGCGCGCG 1072  
QY 294 GluGluLysGlyAlaCysGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 313  
DB 1073 GACGAGCGCTCGGCC-----TCGCGCGCGTGTGTGTGTGTGTGTGTGTGTGTGT 1111  
QY 314 AlaValProAla---LysArgLysAspLeuHisMetLeuValTrpSerSerAlaSer 332  
DB 1112 GCCGTGCCAGCCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1171  
QY 333 ProValSerGluArgAlaAlaValHisValPheGlyAlaGlyGlyAlaAspHisAlaAsp 352  
DB 1172 CCGGTCTCGGAGCTCAGCGCGCTCCCGCGCTTCAACGCGTGTGTGTGTGTGTGTGTGT 1231  
QY 353 ValLeuAlaLys----- 356  
DB 1232 GTCGGCGCGAAGAAATTCGATGTTGTCTCCCGCGGAGATCGCGCAGAACGCGCTCGGC 1291  
QY 357 -----GlyAlaGlnAla-----TyrAspGluTrp 364  
DB 1292 GGCAGAGAGACGAGAACAAAGTGTGACGACGACGACGCGCGGGAAGCGCGGCTTC 1351  
QY 365 Gly---ArgAspAspTyrSerSerArgThrLysAsnGlySerGlyGlyAlaAspLysGly 383

DB 1352 GGCTTCAGATCAGCGCGCAGACGCGCGGAGGAGCGCGGCGGAGAGCAGCAGGG 1411  
QY 384 GlyProThrLeuSerLysLeuGlySerAsn---SerThrAlaGlnLeuTrpProLysAsp 402  
DB 1412 CCGAGCAGCTAACGAGCTGGGTCCAGGTGTCCAGCGCGCGGCTGCGCGGAGGAC 1471  
QY 403 AspGlyGluGlyArgAla-----AlaAlaValAlaMetProProAlaSerValMetThr 420  
DB 1472 GACGCGGTGTGGCGCGGCTCGCGCGCAGACAGATGCGCGCGCGCGAGCTGTATGACC 1531  
QY 421 ArgLeuIleLeuIleMetValTrpArgLysLeuIleArgAsnProAsnThrTrpSerSer 440  
DB 1532 CGCTGATATTAATCATGTGTGTGGCGCACTGATCGCAACCGCAACACGTAATCCAGC 1591  
QY 441 LeuIleGlyValValTrpSerLeuValSerTrpArgTrpGlyIleGluMetProAlaIle 460  
DB 1592 CTATCGCGCTCGCTGTCTCTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1651  
QY 461 IleAlaArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeu 480  
DB 1652 GTGCGCAAGTCCATCGCATCTCTCGACGCGCGGTGTGGGTGTGGGTGTGGGTGT 1711  
QY 481 GlyLeuPheMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAlaAlaIle 500  
DB 1712 GGATGTGTATGCGCTGTGCGCGAGCTCATCGTGTGTGTGTGTGTGTGTGTGTGT 1771  
QY 501 AlaMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaSerIleAlaVal 520  
DB 1772 TCCATGCGCGCTCGCTCTCTCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1831  
QY 521 GlyLeuArgGlyValLeuLeuHisIleAlaIleValGlnAlaAlaLeuProGlnGlyIle 540  
DB 1832 GGCTTCGAGGCGCGCTTTTGGGATCGCATTTGTGACGCGCTCTACCGAAGGATC 1891  
QY 541 ValProPheValPheAlaLysGlyTrpGlyValHisProAspIleLeuSerThr----- 558  
DB 1892 GTGCGCTTCTGT 1951  
QY 559 AlaTrpGly-----ProIleThr 564  
DB 1952 ATTTTGGCATGTGATAGCGCTGCCGATCACC 1984  
RESULT 7  
US-09-938-842A-847  
; Sequence 847, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE OF INVENTION: SAME, AND METHODS OF USE  
; FILE REFERENCE: SRIPI300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 847  
; LENGTH: 1923  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-847  
Alignment Scores: 2.58e-152 Length: 1923  
Pred. No.: 1560.50 Matches: 353  
Score:

Percent Similarity: 63.87% Conservative: 73  
 Best Local Similarity: 52.92% Mismatches: 102  
 Query Match: 53.63% Indels: 139  
 DB: 9 Gaps: 20

US-10-030-884-14 (1-573) x US-09-938-842A-847 (1-1923)

QY 1 MetIleThrAlaLeuAspLeuThrHisValLeuThrAlaValProLeuThrValAla 20  
 DB 1 ATGATCTCATGCGACGAGCTCTACACGGTCTCTACCGCGGTGATTCCTCTCTACGTGGCC 60

QY 21 MetThrLeuAlaThrGlySerValArgTrpArgIlePheThrProAspGlnCysSer 40  
 DB 61 ATGATCTCTCGTTACGGTCTCTCGGTGGTGGAAATCTCTCACCGACCAATGCTCC 120

QY 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
 DB 121 GGAATCAACCGTTTGTGGCATCTCTCGCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180

QY 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80  
 DB 181 ACCAACAATCTTCACGCAATGATCTCTCGTTTCATCGCGCGGACACTCTCCAAAATC 240

QY 81 AlaValLeuAlaLeuAlaLeuAlaSerArgGlyLeuSerSerProArgAlaLeuGly 100  
 DB 241 ATCATGTTGTCACTTTTGTAGTCTATGGCT-----AATTCACCTCGCTCCGGTAGC 291

QY 101 LeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThrLeuValMetGly 120  
 DB 292 CTCGAGTGAGCATCACAACTCTTTCCCTCTCCACATCCACACTCTTGTATGGGG 351

QY 121 IleProLeuLeuArgGlyMetThrGlyAlaSerSerAlaGlyThrLeuMetValGlnVal 140  
 DB 352 ATCTCTCTCTGATCGCATGATGGCAATCTCT---GGTTCCTCATGTCCCAATC 408

QY 141 ValValLeuGlnCysIleIleTrpThrLeuMetLeuPheLeuPheGluTrpArgAla 160  
 DB 409 GTCTGCTCTCCAGTGATCATCTGGTACACGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 468

QY 161 AlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSerIleValSerPhe 180  
 DB 469 GCCAGATCTCTATCATGAGCAGTCTCCCTGAG---ACGGCTGCTCTCCATTGTTCTTTC 525

QY 181 ArgValAspSerAspValValSerLeuAlaArgGlyAsp---ValGluLeuGlnAla 199  
 DB 526 AAAGTCGAATCCGACGCTGCTTTCGTCGACGCCCATGATTTCTTGACACCGATGCAGAG 585

QY 200 ProAspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGlyAspAlaGlyArgVal 219  
 DB 586 ATA-----GGTGACGACGGAGCTT 606

QY 220 ArgValThrValArgLysSerThrSerSerArgSerGluAlaAlaCysSerHisSerHis 239  
 DB 607 CAGCTCACCGTGAGAAATCCACAGCTTCACTG---CGTTCTGCTCTG----- 651

QY 240 SerGlnThrMetGlnProArgValSerAsnLeuSerGlyValGluIleThrSerLeuGln 259  
 DB 652 GGCCCCAATCATGACTCCACGGCGCTCAAAATCTCACGGAGCTGAGATTATAGTCTCAGC 711

QY 260 SerSerArgAsnProThrProArgGlySerPheAsnHisAlaAspPhePheAsnIle 279  
 DB 712 ACC-----ACTCTAGAGCTCTAATTCACCACTCTGATTTTACACATG 759

QY 280 ValGly-----Ala 282  
 DB 760 ATGGGTTTCCCGGTGGTGTCTCTCCAATTTGGTCCGGCGGATATGACTCCGTTCAA 819

QY 283 AlaAlaLysGly-----GlyGlyGlyAlaAlaGly 292  
 DB 820 TCATCTAGAGGTCCAACTCTCGACCTTCAAATCTCGAGAGAAATGCGCCATGCACTCC 879

QY 293 AspGluGlnLysGlyAlaCysGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 310

RESULT 8

US-09-938-842A-847

; Sequence 847, Application US/09938842A

DB 880 TCCCCGAGATTCCGGTATTACCTCGAGGAGGCGCGGTCTTATCCGGCTCCGAATCCA 939  
 QY 310 ----- 310

DB 940 GAGTTCTCTTACCAACCACTCTACCGCCAATAAAGCGTCAATAAAAAACCGAAGAC 999

QY 311 -----GlnAlaValAlaValProAla-----LysArg 319

DB 1000 GTTAATACGATCAGCAGGAGCTCTTCAACGGGCGGCAAGTCAACACGCTAGCGCC 1059

QY 320 LysAspLeuHisMetLeuValTrpSerSerSerAlaSerProValSerGluArgAlaAla 339

DB 1060 AAGGAGCTTCAATGTTCTGCTCGAGCTCAACGGGTCAACCGGTTTTCAGACCGGGGGT 1119

QY 340 ValHisValPheGlyAlaGlyGlyAlaAspHis-----AlaAspValLeu 354

DB 1120 CTTAAGCTTTTTCGCG---GGAGCACTGCAACGATCAAGCGGGAAGTCTGACCAAGT 1176

QY 355 AlaLys----- 356

DB 1177 GCTAAAGAGATCCGTATGTTAGTCCAGATCAATCTCACAAACGGCGAGCAAGCTGTA 1236

QY 357 -----GlyAlaGlnAlaThrAspGluTrpGlyArgAsp 367

DB 1237 GCTCATCCAGCAAGTGGAGATTTCCGAGGAGAAACAATTTAGTTTCCGCCGAAAGAA 1296

QY 368 AspTrpSerSerArgThrLysAsnGlySerGlyGlyAlaAspLysGlyGlyProThrLeu 387

DB 1297 GAAGACGAGAGAGACCAAAAGACCGCGAGATGTT-----CTA 1335

QY 388 SerLysLeuGlySerAsnSerThrAlaGlnLeuTrpProLysAspAspGlyGlyArg 407

DB 1336 AACAACTTCTCTCCAAATTTCCACGGCGGCTCAATCAAG---ACAGCTTAGAGGA 1392

QY 408 AlaAlaAlaValAla-----MetProAlaSerValMetThrArgLeuLeu 424

DB 1393 GCCGAGCAAGTCAACGAAAAATATGCTCCCGCGAGTGTGATGACAAAGCTGATACG 1452

QY 425 IleMetValTrpArgLysLeuIleArgAsnProAsnThrTrpSerSerLeuIleGlyVal 444

DB 1453 ATAATGTTTGGAGGAACTCATCAGAAACCCAAACACTTACTCTACTCTCATTTGACCT 1512

QY 445 ValTrpSerLeuValSerTrpArgTrpGlyIleGluMetProAlaIleIleAlaArgSer 464

DB 1513 ATTTGGCTCTCTCGCTTTCGGTGGCAGCTGCAATGCCAAATCATTCACCAATCT 1572

QY 465 IleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeuGlyLeuPheMet 484

DB 1573 ATCTCCATCTATCTGATGCTGCTCTTGGAAATGCAATGTTTGTGGGTGTTTCATG 1632

QY 485 AlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAlaIleAlaMetGlyVal 504

DB 1633 CGGTTCGCAACCCAAATTAATCGTTTGGGAATTCAGTGGCAACGTTTGCATGGCGGTT 1692

QY 505 ArgPheValAlaGlyProAlaValMetAlaAlaAlaSerIleAlaValGlyLeuArgGly 524

DB 1693 AGGTTCTTACGGGTCCGGCGGTGATGGCGGTGCTGTATAGCCATCGGATTCAGTGGT 1752

QY 525 ValLeuLeuHisIleAlaIleValGlnAlaAlaLeuProGlnGlyIleValProPheVal 544

DB 1753 GATTACTCGGTGTCGTATAGTTTACGCGCATTACTCAGGAATTTGCCCTTTGTG 1812

QY 545 PheAlaLysGluTrpGlyValHisProAspIleLeuSerThrAla-----TyrGly--- 561

DB 1813 TTTCCGAGGAGTACATGTTTCATCTCTGCTATTATTAAGTACAGGGTAAATTTTGAATG 1872

QY 562 -----ProIleThr 564

DB 1873 CTTATAGCGTTCCGATCAGC 1893



```
QY 465 IleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeuGlyLeuPheMet 484
Db 1573 ATCTCCATCTATCTGATGCTGGTCTTGGAAATGCAATGTTAGTTGGGGTGTTCATG 1632
QY 485 AlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAlaIleAlaMetGlyVal 504
Db 1633 CGCTTGCAACCAATTAATCGCTTGTGGAAATTCAGTGGCAACGTTTGGCATGGCGGTT 1692
QY 505 ArgPheValAlaGlyProAlaValMetAlaAlaSerIleAlaValGlyLeuArgGly 524
Db 1693 AGGTTCCCTTACGGTCCGGCGTGTATGCGGTTGCTGTATAGCCATCGGATACGTGGT 1752
QY 525 ValLeuLeuHisIleAlaIleValGlnAlaAlaLeuProGlnGlyIleValProPheVal 544
Db 1753 GATTTACTCGCTGCTGCTATAGTTTCAAGCCGCAITACCTCAAGAAATGTGCGCCCTTGTG 1812
QY 545 PheAlaLysGlyIleGlyValHisProAspIleLeuSerThrAla-----TyrGly--- 561
Db 1813 TTTCGAAGGAGTACAATGTTCATCTCTGCTATTTTAAGTACAGGGGTAAATTTTGAATG 1872
QY 562 -----ProIleThr 564
Db 1873 CTTATAGCGCTTCCGATCAAG 1893
RESULT 9
US-09-887-576-628
; Sequence 628, Application US/09887576
; Patent No. US20020144047A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, P.
; APPLICANT: Brown, D.
; APPLICANT: Chang, H.
; APPLICANT: Zhu, T.
; APPLICANT: Han, B.
; APPLICANT: Wang, X.
; APPLICANT: Cooper, Bret
; TITLE OF INVENTION: Promoters for regulation of plant expression
; FILE REFERENCE: 1360.001US1
; CURRENT APPLICATION NUMBER: US/09/887,576
; CURRENT FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/213,848
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/214,087
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/258,692
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 875
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 628
; LENGTH: 2222
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-887-576-628
Alignment Scores:
Pred. No.: 2,386-150 Length: 2222
Score: 1542.50 Matches: 349
Percent Similarity: 62.58% Conservative: 59
Best Local Similarity: 53.53% Mismatches: 111
Query Match: 53.01% Indels: 133
DB: 9 Gaps: 17
US-10-030-884-14 (1-573) x US-09-887-576-628 (1-2222)
QY 1 MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValProLeuTyrValAla 20
Db 1 ATGATTACGGCGACGACTGTACACCGTCTTCACCGCGGTGGTACCACTTTACGTAGCT 60
QY 21 MetThrLeuAlaTyrGlySerValArgTrpTrpArgIlePheThrProAspGlnCysSer 40
Db 61 ATGATTCTCGCTACGGATCCGTACAGTGGTGAAGATATCTCACCAGACAGTGTCTCC 120
QY 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60
```

```
Db 121 GGCATCAACCGCTTCGTGCTATCTTCGCGGTCCTCTCCTCTCCCTCCACTTCATCTCC 180
QY 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80
Db 181 ACCAACGATCCTTACGCCATGAATTCGCTTCGCGCCGACACGCTTCCAAAATC 240
QY 81 AlaValLeuAlaLeuAlaLeuAlaSerArgGlyLeuSerSerProArgAlaLeuGly 100
Db 241 ATCATGCTCGTCTTACTTCTCTATGGGCT-----AACCTAACCAAGAACGGTAGC 291
QY 101 LeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThrLeuValMetGly 120
Db 292 TTGGAGTGATGATCACAACTTCTCTCAGCATCTCCCAACACACTTGTTCATGGGG 351
QY 121 IleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeuMetValGlnVal 140
Db 352 ATCCCTCTGTGATCGCCATGTACGGA---ACCTACGCAAGGTCTCTAATGGTCCAAAGTC 408
QY 141 ValValLeuGlnCysIleIleTrpTyrThrLeuMetLeuPheLeuPheIleTyrArgAla 160
Db 409 GTTGTCTTCAGTGATCATTTTGGTACACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 468
QY 161 AlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSerIleValSerPhe 180
Db 469 GCTAAGCTTCTTATTATGGAGCAGTTCGCGAG---ACTGGTGCCTTATTGTTCGTTT 525
QY 181 ArgValAspSerAspValValSerLeuAlaArgGlyAsp---ValGlnLeuGlnAlaGln 199
Db 526 AAAGTTGAATCCGACGTCGTTTCTCTCGACGCTCATGATTTTCTTGAGACGATGCTGAG 585
QY 200 ProAspGlyValAlaGlyAlaGlyAlaValSerArgGlyGlyAspAlaGlyArgVal 219
Db 586 ATA-----GGAAACGACGGAGCTT 606
QY 220 ArgValThrValArgLysSerThrSerSerArgSerGluAlaAlaCysSerHisSerHis 239
Db 607 CATGTTTACCGTGAGAAATCAAAACGATCGAGA-----CGG 642
QY 240 SerGlnThrMetGlnProArgValSerAsnLeuSerGlyValGluIleTyrSerLeuGln 259
Db 643 TCACGTGATGACTCCACGGCTTCGAATCTTACCGAGCTGAGATCTATAGTCTTAGT 702
QY 260 SerSerArgAsnProThrProArgLysSerSerPheAsnHisAlaAspPhePheAsnIle 279
Db 703 TCG-----ACTCCGAGAGGTTCTAACTTTAACCAITCTGATTTTACTCTGTT 750
QY 280 ValGlyAlaAlaAlaLysGlyGlyGlyAlaAlaGlyAsp----- 293
Db 751 ATGGGGTTTCCC-----GGCGGAGGCTTTTCAATTTTGTGTCGGCGGATTTG 798
QY 294 -----GluGluLysGly 297
Db 799 TACTCCGTTCAATCTTCTGTCGTCGACTCCACGCGCTTCGAACTTCGAAGAGAACAC 858
QY 298 AlaCysGlyGlyGly-----GlyGlyGlyHisSerProGlnProGlnAlaValAlaVal 315
Db 859 GCCGTTAAATATGGAATTTACAATAACACTAACAAGTTCTGTTCGGCGCGCGGTTCTG 918
QY 316 ProAlaLysArg----- 319
Db 919 CCGGCTCCGACCCCGAGTTTCAACCGGTACGGGTGTTTCAACTAAACCGAATAAAATT 978
QY 320 -----Lys 320
Db 979 CCTAAAGAAACCAACAGCAACTGCAAGAGAAAGATAGCAAGCGAGCCATGACGTAAG 1038
QY 321 AspIleHisMetLeuValTrpSerSerSerAlaSerProValSerGluArgAlaAlaVal 340
Db 1039 GAGCTTCACATGTTTGTGGAGCTCAGCGCTTCTCCGCTCTCCGAC----- 1086
QY 341 HisValPheGlyAlaGlyAlaAspHisAlaAspValLeuAlaLysGlyAlaGlnAla 360
```



1087 ---GTGTTTGGCGAGGTGACGCGACACACGTG-----GCAACGGAA 1125  
Db  
361 TyrAspGluTyrGlyArgAspTyr-----SerSerArgThrLys 374  
Qy  
1126 CAATCTGAACAAGGTGCTAAGAGATTGGATGGTGTCTGATCAACCTCGAAGAGT 1185  
Db  
375 AsnGlySerGlyGlyAlaAsp----- 381  
Qy  
1186 AATGCTAGAGGTGGTGTGATGATATCGCGGTCTTGATGTGAGAGGAGAAAGAG 1245  
Db  
382 -----LysGlyGlyProThrLeuSerLysLeuGlySerAsnSerThrAlaGlnLeuTyr 399  
Qy  
1246 ATAGAGAAAGCTACAGCGGTGTAATAATGGGGTCTTAATTCACCGCGGAGTAGAG 1305  
Db  
400 ProLysAspAspGlyGlyAlaAlaAlaValMetProProAlaSerValMet 419  
Qy  
1306 GCGGCTGGTGGAGATGCGCGCGCAACACGACACATATGCCCGCCGACAAAGTGTGATG 1365  
Db  
420 ThrArgLeuLeuLeuLeuMetValTyrArgLysLeuLeuLeuAsnProAsnThrTyrSer 439  
Qy  
1366 ACACGACTGATATTGATAATGTGTGGAGAAAGCTGATCAGAAACCCAAACACGACTCC 1425  
Db  
440 SerLeuLeuGlyValValTyrSerLeuValSerTyrArgTyrGlyLysLeuMetProAla 459  
Qy  
1426 AGTCTAATCGGTCTCATATGGGCTCTTGTGTTACCGGTGGCATGTGGCTATGCCCAA 1485  
Db  
460 IleLeuAlaArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPheSer 479  
Qy  
1486 ATATTACAACAATCCATCTCCATCTCTCAGATGCTGGTCTTTGGAATGGCTATGTTACG 1545  
Db  
480 LeuGlyLeuPheMetAlaLeuGlnProArgIleLeuAlaCysGlyAsnLysLeuAlaAla 499  
Qy  
1546 TTAGGTTTATTCATGGCATTTCAACCCAAATCATTGTTGGGACTCTGTGGACACG 1605  
Db  
500 IleAlaMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaSerIleAla 519  
Qy  
1606 TTGGCATGGCGTCAATTTATAACCGGTCCGGCCATCATGGCTGTGGATTGCC 1665  
Db  
520 ValGlyLeuArgGlyValLeuLeuHisIleAlaIleValGlnAlaAlaLeuProGlnGly 539  
Qy  
1666 ATGGGCTTACAGCGGACCTTCTCCGTATAGCATCTGTTACGGCTGCGTGGCTCAAGA 1725  
Db  
540 IleValProPheValPheAlaLysGluTyrGlyValHisProAspIleLeuSerThrAla 559  
Qy  
1726 ATAGTTCCGTTGTGTTGCAAAAGAGTACAATGTGCATCCACCATTTAAGCACTGGG 1785  
Db  
560 -----TyrGly-----ProIleThr 564  
Qy  
1786 GTCATATTGGAAATGTTAATAGCCTTACCTATAACT 1821  
Db

## RESULT 10

US-10-425-114-1349  
; Sequence 1349, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 1349  
; LENGTH: 2276  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700151120\_FLI

## US-10-425-114-1349

Alignment Scores:  
Pred. No.: 2,78e-150 Length: 2276  
Score: 1542.00 Matches: 357  
Percent Similarity: 65.61% Conservative: 57  
Best Local Similarity: 56.58% Mismatches: 135  
Query Match: 52.99% Indels: 82  
DB: 12 Gaps: 19

## US-10-030-884-14 (1-573) x US-10-425-114-1349 (1-2276)

Qy 1 MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValProLeuTyrValAla 20  
Db 115 ATGATCACCAGCGCGGACATCTACAGCTTCTTGGCGCATCGTGGCGGTGACGTGGCC 174  
Qy 21 MetThrLeuAlaTyrGlySerValArgTyrPheArgIlePheThrProAspGlnCysSer 40  
Db 175 ATGTTCTGGCGTACGGTTCGGTGGTGGCCATCTTACCCCGGACCACTGCTCTCC 234  
Qy 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
Db 235 GGCATCAACCGCTTCGTGGCGGTCTTGGCGGTGGCCGTCTCTCTTCCCTTCCATCTCC 294  
Qy 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80  
Db 295 AGCTGGACCCCTACGCGATGAGTACCGGTCTTGGCGCGCGCATCTCGTGCAGAAAGCTG 354  
Qy 81 AlaValLeuAlaLeuLeuAlaLeuAlaSerArgGlyLeuSerSerProArg----- 97  
Db 355 GTCATCTGGCGCGTGGCGGTGGCGACACGCTCTCTCCGCTACCGCGGCGGCGG 414  
Qy 98 ---AlaLeuGlyLeuAspTyrPheIleThrLeuPheSerLeuSerThrLeuProAsnThr 116  
Db 415 GCGGCGTGTCTGGTGGACCATCACGCTCTTCTCGTGTCCAGCTGCTCCCAACACG 474  
Qy 117 LeuValMetGlyIleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeu 136  
Db 475 CTGCTGATGGCATCTCCGCTGCTCCGCGCATATACGGCGACTTCTCG---GGCAACCTC 531  
Qy 137 MetValGlnValValLeuGlnCysIleIleTyrThrLeuMetLeuPheLeuPhe 156  
Db 532 ATGTGTCAGATCTGTGTGTCTGAGAGGTCATCTGTACAGCTCATGTCTCTCTCTTC 591  
Qy 157 GluTyrArgAlaAlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaSer 176  
Db 592 GAGTACCGCGCGCAAGCGCTCATCTCCGAGCAGTTCCCGCCGCGACGCTCGCGCGCAGC 651  
Qy 177 IleValSerPheArgValAspSerAspValValSerLeuAlaArgGlyAspValGluLeu 196  
Db 652 ATCGCTCTCTTCAAGGTGAGTCCGACGTCTCTCTCTC---AACGGCGCGAGGCGCTG 708  
Qy 197 GluAlaGluProAspGlyValAlaGlyAlaValSerSerArgGlyGlyAspAla 216  
Db 709 CAGCGCGAC-----GCCGAGGTCCGCGACGAC 735  
Qy 217 GlyArgValArgValThrValArgLysSerThrSer----- 228  
Db 736 GCGCGCTTCCACGCTGTCATCCGCGCTCTCCGCTCCCGCTCCACCGCGCGGCGGCGAC 795  
Qy 229 -----SerArgSerGluAlaAlaCys---SerHisSerHisSerGlnThrMetGlnPro 245  
Db 796 GCGCGCGCGCTCCGCGCGCGGTACCGCCCTACGCGCGCTCTCGGCGCATGACCCCG 855  
Qy 246 ArgValSerAsnLeuSerGlyValGluIleTyrSerLeuGlnSerSerArgAsnProThr 265  
Db 856 CGCGCTCCCACTCACCGCGGTGAGATCTACTGCTGAGAGCTCGCGGAGCCACCG 915  
Qy 266 ProArgGlySerSerPheAsnHisAlaAspPheAsnIleValGlyAla----- 282  
Db 916 CCCCGCGCTCCAGCTTCAACCATGCTTCTACCGCATGTTCACCGGAGCAAGATG 975  
Qy 283 -----AlaAlaLysGlyGlyGlyAlaAlaGly----AspGluGlu----- 295



APPLICANT: Kreps, Joel  
APPLICANT: Wang, Xun  
APPLICANT: Zhu, Tong  
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
TITLE OF INVENTION: SAME, AND METHODS OF USE  
FILE REFERENCE: SCRIPI300-3  
CURRENT APPLICATION NUMBER: US/09/938,842A  
CURRENT FILING DATE: 2001-08-24  
PRIOR APPLICATION NUMBER: US 60/227,866  
PRIOR FILING DATE: 2000-08-24  
PRIOR APPLICATION NUMBER: US 60/264,647  
PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: US 60/300,111  
PRIOR FILING DATE: 2001-06-22  
NUMBER OF SEQ ID NOS: 5379  
SEQ ID NO 848  
LENGTH: 1860  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
US-09-938-842A-848

Alignment Scores:  
Pred. No.: 9,64e-149 Length: 1860  
Score: 1526.00 Matches: 346  
Percent Similarity: 64.59% Conservative: 68  
Best Local Similarity: 53.98% Mismatches: 119  
Query Match: 52.44% Indels: 108  
DB: Gaps: 19

US-10-030-884-14 (1-573) x US-09-938-842A-848 (1-1860)

Qy 1 MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValValProLeuTyrValAla 20  
Db 1 ATGATCACATGGCAGCAGCTTACACCGTCTTACCGCGGTGATACCACTCTACGTAGCC 60  
Qy 21 MetThrLeuAlaTyrGlySerValArgTyrTrpArgIlePheThrProAspGlnCysSer 40  
Db 61 ATGATCCTCGCTTACCGCTCAGTCCGGTGGTGAATAATCTTCTCACGACCACTTCTTC 120  
Qy 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
Db 121 GGCATAAACCGCTTCGTGCTATTTTCGCGCTCTCTCTCTCTCTCTCTCTCTCTCT 180  
Qy 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnVal 80  
Db 181 TCAACAACATCTTACGCAATGAATCTCCGATTCATCGCAGCGGATACATCCAAAACCTA 240  
Qy 81 AlaValLeuAlaLeuAlaLeuAlaLeuAlaSerArgGlyLeuSerSerProArgAlaLeuGly 100  
Db 241 ATCATGCTCACTCTCCATATCACTGGGCT-----AACTTCACTCGCTCCGGTAGT 291  
Qy 101 LeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThrLeuValMetGly 120  
Db 292 CTTGAATGAGGAGCATCAATCTTCTCTTTTCAACACTTCCCAACACTCTCTCTTATGGGA 351  
Qy 121 IleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeuMetValGlnVal 140  
Db 352 ATACCTCTCTTGATCGCCATGTCGAGAGATATATCC---GGCTCACTCATGGTTCAGATC 408  
Qy 141 ValValLeuGlnCysIleIleTyrThrLeuMetLeuPheLeuPheGlyTyrArgAla 160  
Db 409 GTTGTCTTCTCAGTGATATAATCTGTACACGCTTCTCTCTCTCTCTCTCTCTCTCTCT 468  
Qy 161 AlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSerIleValSerPhe 180  
Db 469 GCTAAGATCTTGATCATGGAACAGTTTCCAGAG---ACTGGTCTTCTCGATGATCTGTT 525  
Qy 181 ArgValAspSerAspValValSerLeuAlaArgGlyAsp---ValGluLeuGluAlaGlu 199  
Db 526 AAAGTTGAGTCCGATGTTGTTCTTTAGATGGACATGATTTCTTTGAAACTGATGCTCAA 585  
Qy 200 ProAspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGlyAspAlaGlyArgVal 219  
Db |||||

976 GCCAGCCGCTGGGCGAGCG 1035  
Qy 296 ---LysGlyAlaCysGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 311  
Db 1036 AACAGATTCTGGGAAGGGTGGGACCGCGCGGTACCGCGCGCGCGCGCGCGCGCGCGCG 1095  
Qy 312 AlaValAlaValProAlaLysArg-----LysAsp 321  
Db 1096 ATGATGCCACCGCCACGAGAGAGAGTGGCGGCTCCAACTCCAACTCCAACTCCAACTCC 1155  
Qy 322 LeuHisMetLeuValTrpSerSerSerSerAlaSerProValSerGlu-----ArgAla 338  
Db 1156 CTGCACATGTTGTTATGAGTCCAGCGCGTGGCGCGTGGCGCGTGGCGCGTGGCGCGTGG 1215  
Qy 339 AlaValHisValPheGlyAlaGlyGlyAlaAspHisAla-----AspVal 353  
Db 1216 GCCGTCAAC-----CAGCGCGCTCCACCGACTCGCGCGCGCGCGCGCGCGCGCG 1269  
Qy 354 LeuAlaLysGlyAlaGlnAlaTyrAspGluTyrGlyArgAspAspTyrSerSerArgThr 373  
Db 1270 CCGGTCCGACGGCGCGCCACCCCAAGAGCGTGAAGTGGACTGTCCCGCGCGCGCGCGCG 1329  
Qy 374 LysAsnGlySerGlyGlyAlaAsp-----LysGlyGlyProThr 386  
Db 1330 GACCCCG 1389  
Qy 387 -----LeuSerLysLeuGlySerAsnSerThrAlaGlnLeuTyrProLysAspAspGly 404  
Db 1390 GGGCTGGCGCGCAAGTTCCTCCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCG 1449  
Qy 405 GluGly-----ArgAlaAlaValAlaMetProProAlaSerValMetThr 420  
Db 1450 GCCGACGGCG 1509  
Qy 421 ArgLeuIleLeuIleMetValTrpArgLysLeuIleArgAsnProAsnThrTyrSerSer 440  
Db 1510 CGCTCATCTCATCTGTTGGAGAGCTGATCAGCAACCCCAACCACTACTCTCCAGC 1569  
Qy 441 LeuIleGlyValValTrpSerLeuValSerTyrArgTyrGlyIleGluMetProAlaIle 460  
Db 1570 CTCATCGGCTCGTCTGTCGCTCGTCTGTTTAAACATCCAGATCCGCTCGGAT 1629  
Qy 461 IleAlaArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeu 480  
Db 1630 ATAAGGATTCGATATCTATATATCCGACGCGAGGCTAGGATGCGCATGTTCACTTA 1689  
Qy 481 GlyLeuPheMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAlaIle 500  
Db 1690 GGTCTCTCATGCTCTGCAACCGAAGCTCATCTCTCTGCGTAAAGAGGTGGCGACGTT 1749  
Qy 501 AlaMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaAlaSerIleAlaVal 520  
Db 1750 GCCATGGCGGTGAGGTTTGTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1809  
Qy 521 GlyLeuArgGlyValLeuLeuHisIleAlaIleValGlnAlaAlaLeuProGlnGlyIle 540  
Db 1810 GGCTCCGGGAGTGTCTCTGACGCGTGCATCTGTTCAAGCTGCACTTCCCAAGGCGATC 1869  
Qy 541 ValProPheValPheAlaLysGluTyrGlyValHisProAspIleLeuSerThrAla--- 559  
Db 1870 GTCCCTCTGTTGTTGCCAAGGAGTACAACTGCCATCTCTCAAAATCTTAGCACCGGTT 1929  
Qy 560 ---TyrGly-----ProfileThr 564  
Db 1930 ATATTCGGGATGCTGATCGCGCTTCGATCAGC 1962

RESULT 11  
US-09-938-842A-848  
Sequence 848, Application US/09938842A  
Patent No. US20020160378A1  
GENERAL INFORMATION:  
APPLICANT: Harper, Jeff

```

586 ATC-----GGTGACGATGGTAAGCTT 606
220 ArgValThrValArgLysSerThrSerArgSerGluAlaAalaCysSerHisSerHis 239
607 CARGTTACGGTGAGAAATCAACGGTCTTCGG-----AGATCTTTTACCGTGGTGT 660
240 SerGlnThrMetGlnProArgValSerAsnLeuSerGlyValGluLeuTyrSerLeuGln 259
661 GGTAATAATAGTACTCTCGTCGTCTAATCTACCGGAGCTGAGATTATAGTCTT-- 717
260 SerSerArgAsnProThrProArgGlySerSerPheAsnHisAlaAspPheAsnIle 279
718 -----AACACTACTCAAGAGAGTCTAACTTCAATCATCTGATTTTACTCTATG 768
280 ValGly-----Ala 282
769 ATGGGGTTCCCGGTGGCGGCTTTTCGAACCTTTGGTCCGGCGGATATGACTCTGTCAG 828
283 AlaAlaLysGly-----GlyGlyAlaAlaGly 292
829 TCCTCGAGAGGACTACTCTAGACCTTCGAATTTTGGAGGAGTGTGTCTATGCTTCT 888
293 AspGluGluLysAlaCysGlyGlyGlyGlyGlyGlyHisSerProGlnProGln--- 311
889 TCGCCAAGATTCGGGTATTACCCGGGAGGAGCTCCCGGTCTGCTACCCAGCTCTTAATCCG 948
312 -----AlaValAlaValProAlaLysArg----- 319
949 GAGTTTTTACCGGTAATAAAACCGGTAGTAAGCTCCAAAGAGAAATCATCATGTTA 1008
320 -----LysAspLeuHisMetLeuValTrpSerSerAla 331
1009 GGAATATCAATAGTAAATGATGCTAAGGAGCTTCATATGTTTGTGTGGGATCCACCGA 1068
332 SerProValSerGluArgAlaAlaValHisVal---PheGlyAla----- 345
1069 TCACCGGTTTCGACCGAGCTGCTCTTCAAGTTGATAATGGAGCCAATCAACAAGTCGA 1128
346 -----GlyGlyAlaAspHisAlaAspValLeuAla----- 355
1129 AAATCCGATCAAGCGGTGCAAAAGAGATTCGAATGTTGATCTCTGATCATACTCAAAAT 1188
356 -----LysGlyAlaGlnAlaTyrAspGluTyrGlyArgAspAspTyrSerSerArg 372
1189 GGTGAAACAAAGCTGGTCCGATGAACGGGACTATGGCGGGAAGAGAGTCGGAGAGG 1248
373 ThrLysAsnGlySerGlyGlyAlaAspLysGlyGlyProThrLeuSerLysLeuGlySer 392
1249 GTAAAGGAGAGTGCCTAACGGA-----CTACACAAGCTTCGGTGT 1287
393 AsnSerThrAlaGlnLeuTyrProLysAspAspGlyGlu---GlyArgAlaAlaVal 411
1288 AATCCACAGCAGCTTAACCTTAAGAGCTATAGAACGGGTGAACACTGTACCGGTA 1347
412 Ala---MetProAlaSerValMetThrArgLeuLeuLeuMetValTrpArgLys 430
1348 AAACATATGCCACAGCGAGTGTGATGACTCGGCTGATATTCATATGTTGGTGGGAAA 1407
431 LeuLeuArgAsnProAsnThrTyrSerSerLeuLeuGlyValValTrpSerLeuValSer 450
1408 CTATAGAAACCAACCACTTACTTAGTCTCATTTGGTCTCATTTGGGCTTTTGTGCT 1467
451 TyrArgTrpGlyIleGluMetProAlaIleIleAlaArgSerIleSerIleLeuSerAsp 470
1468 TTCAGGTGGGATGTGGCAATGCTCAAAATTAATCAACAATCAATCTCAATCTTCTGAT 1527
471 AlaGlyLeuGlyMetAlaMetPheSerLeuGlyLeuPheMetAlaLeuGlnProArgIle 490
1528 GCTGGTCTCGTATGGCAATGTTAGTGTGGGTTGTTCATGGCATTCGACCGCAAAATTA 1587
491 IleAlaCysGlyAsnLysLeuAlaAlaIleAlaMetGlyValArgPheValAlaGlyPro 510
1588 ATTGCTTGGGAATTCGACGGGAGCTTTTGGGATGGCGTGAGATTCTTTACTGGACCA 1647

Qy 511 AlaValMetAlaAlaAlaSerIleAlaValGlyLeuArgGlyValLeuLeuHisIleAla 530
Db 1648 GCGTAATAGCGCGTGGCAGCAATGGCTATTGGATTAGCTGAGACCTATTTCGTGGGCC 1707
Qy 531 IleValGlnAlaAlaLeuProGlnGlyIleValProPheValPheAlaLysGlyTyrGly 550
Db 1708 ATTGTTCAAGCTGCTCAAGGATCGTCCGTTTGTGTTTGCATAAAGAGATATAAT 1767
Qy 551 ValHisProAspIleLeuSerThrAla-----TyrGly-----ProIle 563
Db 1768 GTTCATCCCGCAATCTTGAGTACAGGGTAAATTTTGGAAATGCTTATCGCACTACCGATT 1827
Qy 564 Thr 564
Db 1828 ACA 1830

RESULT 12
US-09-938-842A-848
; Sequence 848, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 848
; LENGTH: 1860
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-848

Alignment Scores:
Pred. No.: 9,64e-149 Length: 1860
Score: 1526.00 Matches: 346
Percent Similarity: 64.59% Conservative: 68
Best Local Similarity: 53.98% Mismatches: 119
Query Match: 52.44% Indels: 108
DB: 11 Gaps: 19

US-10-030-884-14 (1-573) x US-09-938-842A-848 (1-1860)
Qy 1 MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValProLeuTyrValAla 20
Db 1 ATGATCATATGGCAGCACCTCTACCGTCTCTACCGCGGTGATACCACTCTACGTAGCC 60
Qy 21 MetThrLeuAlaTyrGlySerValArgTyrTrpArgIlePheThrProAspGlnCysSer 40
Db 61 ATGATCTTCGCTTACCGCTCAGTCCGGTGGTGGAAATCTTCTCACCAGACCAATGTTCC 120
Qy 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60
Db 121 GGCATAAACCGCTTCGTCGCTATTTCGCGCTCCCTCTCTCTCTTTCCTCATCTCC 180
Qy 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80
Db 181 TCAAACAATCTTACCGCAATGAATCTCCGATTTCATCGACCGCATACACTCCAAACTA 240
Qy 81 AlaValLeuAlaLeuLeuAlaLeuAlaSerArgGlyLeuSerSerProArgAlaLeuGly 100
Db 241 ATCATGCTCACTCTCTCTTAATCATCTGGGCT-----AACTTCACTCGCTCCGGT 291
```



DB: 9 Gaps: 15  
US-10-030-884-14 (1-573) x US-09-938-842A-1305 (1-1983)  
QY 1 MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValValProLeuTyrValAla 20  
DB 1 ATGATCACCAGGAAAGACATGACGATGTTTAGCGGTATGTCGCGCTATACGTTGCT 60  
QY 21 MetThrLeuAlaTyrGlySerValArgTyrTyrTyrTyrTyrTyrTyrTyrTyrTyr 40  
DB 61 ATGATATTAGCTATGTTTCGTTACGTTGTTGGGGATATTCACACCGGACCAATGTTCC 120  
QY 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
DB 121 GGTATAACCGGTTCTGTCGGGTTTTCGGGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCC 180  
QY 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80  
DB 181 TCCATGATCCCTATGCAATGAATACCACTTCTCTGCTGCTGATCTCTTCAGAAATGC 240  
QY 81 AlaValLeuAlaLeuLeu-----AlaLeuAlaSerArgGlyLeuSerSerPro 96  
DB 241 GTTATCTCGCGCACTCTTCTTTGGCAGGGGTTTAGCGCGAGAGGA-----288  
QY 97 ArgAlaLeuGlyLeuAspTyrPheIleThrLeuPheSerLeuSerThrLeuProAsnThr 116  
DB 289 -----AGCCTAGATGGATGATACGCTCTTTTCACTATCAACACTGCTTAAACAG 339  
QY 117 LeuValMetGlyIleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeu 136  
DB 340 TTGGTAATGGGAATCCATGTTAGGCGATGTACGGAGACTTCTCC---GGTAACCTA 396  
QY 137 MetValGlnValValValLeuGlnCysIleIleTyrTyrThrLeuMetLeuPheLeuPhe 156  
DB 397 ATGGTTCAGATCGTGTGCTTCCAGACATCATGTGTATACATTAATGCTCTCTCTGTT 456  
QY 157 GluTyrArgAlaAlaArgAlaLeuLeuAspGlnPheProAspGlyAlaAlaAlaSer 176  
DB 457 GAGTTCGTTGGGTAAGCTTCTCATCTCGAGCAGTTCCTCGAG---ACGCTGTTCA 513  
QY 177 IleValSerPheArgValAspSerAspValSerLeu---AlaArgGlyAspValGlu 195  
DB 514 ATTACTTCTCTCAGAGTTGACTGTGATGTTATCTCTCTTAATGGCGGTGAACCTCCAG 573  
QY 196 LeuGluAlaGlnProAspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGlyAsp 215  
DB 574 ACCGATCGCGAGATA-----GGAGAC 594  
QY 216 AlaGlyArgValArgValThrValArgLysSerThrSerSerArgSerGluAlaAlaCys 235  
DB 595 GACGGAAGCTACACGCTGTGTGTTCAAGATCAAGTGGCGCTCATCATGATCTCTTCA 654  
QY 236 -----SerHisSer-----HisSerGlnThrMetGlnProArgValSerAsn 249  
DB 655 TTCACAAATCTCACGGCGAGACTTAACTCTCTCATGATTAACCGCGAGCTTCAAT 714  
QY 250 LeuSerGlyValGluIleTyrSerLeuGlnSerSerArgAsnProThrProArgGlySer 269  
DB 715 CTCACCGGCTAGAGATTTACTCCGTTCAATCGTCACGAGAGCCGACGCGGAGACTTCT 774  
QY 270 SerPheAsnHisAlaAspPhe-----277  
DB 775 AGCTTTAATCAGACAGATTCTACGCAATGTTTAAACGAAGCAAGCTCCAAAGCCCTCGT 834  
QY 278 -----AsnIleValGlyAlaAlaAlaLysGlyGly-----288  
DB 835 CACGGTTACACTAATAGCTACGCGCGCTGGAGCTGGTCCAGGTGGAGATGTTTACTCA 894  
QY 288 -----288  
DB 895 CTTTCACTTTCTTAAGGCGGTGACGCGGAGAACGTCAAATTTTGTATGAGGAGTTATGAG 954  
QY 288 -----288

DB 955 ACGCGAAGAAAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1014  
QY 289 -----GlyAlaAla 291  
DB 1015 AGTGTTCGCTGTCATCCACCGCGAACCACCAATTTTACGGGGTCAACGAGTGGAGCAAGT 1074  
QY 292 GlyAspGluGlyLysGlyAlaCysGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 311  
DB 1075 GGAGTCAAGAAAGAAAGAGTGGTGGCGGAGGAGCGGTGGC-----1116  
QY 312 AlaValAlaValProAlaLysArgLysAspLeuHisMetLeuValTyrSerSerSerAla 331  
DB 1117 GGAAGTAGGATAGGAGGACAAACCAAGGATGATCAATGTTCTGTGGAGTTCGAGTCT 1176  
QY 332 SerProValSerGluArgAlaAlaValHisValPheGlyAlaGlyGlyAla-----348  
DB 1177 TCTCCGTTGTCGAGGACCAACGCGAAGATGCTATGACGAGAGGTTCTTCCACCGATGTA 1236  
QY 349 -----AspHisAlaAspValLeuAlaLysGlyAlaGln 359  
DB 1237 TCCACCGACCTAAAGTTTCTATTCCTCTCCACGACCAACCTCGTACTTAAAGCGATGCAG 1296  
QY 360 AlaTyrAspGluTyr-----GlyArgAspAspTyrSerSerArgThrLysAsnGly 376  
DB 1297 AATCTGATGAGAACATGTCACCGGAGAAAGGCGATGTGGAATGGACCCAGACGCT 1356  
QY 377 SerGlyGlyAlaAspLysGlyGlyProThrLeuSerLysLeuGlySerAsnSerThrAla 396  
DB 1357 AATAACGGGGGAAG-----TCACCTTACATGGCGCAAAAGGTAGCGAC-----1401  
QY 397 GlnLeuTyrProLysAspAspGlyGlyGlyArgAlaAlaValAlaMetProProAla 416  
DB 1402 -----GTGGAAGACGGCGGTCCCGGTCTAGGAAACAGCAGATGCCCGCGCG 1449  
QY 417 SerValMetThrArgLeuIleLeuIleMetValTyrArgLysLeuIleArgAsnProAsn 436  
DB 1450 AGTGTGATGACGAGACTAATTCGTATAATGGTTTGGAGAAACTCATTCGAAACCTTAAC 1509  
QY 437 ThrTyrSerSerLeuIleGlyValValTyrSerLeuValSerTyrArgTyrGlyIleGlu 456  
DB 1510 ACTTACTCTAGTCTTCTTGGCCCTTGTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1569  
QY 457 MetProAlaIleAlaArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAla 476  
DB 1570 ATGCCACGATATGAGTGGATCGATTTCGATATATTCGATCTGCTTGAATGGCT 1629  
QY 477 MetPheSerLeuGlyLeuPheMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLys 496  
DB 1630 ATGTTAGTCTTGTCTATTTATGGCATTCACCAAGATTTATCGTCGCGGAAATCA 1689  
QY 497 LeuAlaAlaIleAlaMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaAla 516  
DB 1690 GTAGCAGGTTTGGATGGCGGTAAAGTTCTTGACTGGACCGCTGATCGCAGCCACC 1749  
QY 517 SerIleAlaValGlyLeuArgGlyValLeuLeuHisIleAlaIleValGln 533  
DB 1750 TCAATGCAATTTGTTATTCGAGGTGATCTCTCCATATCGCCATCGTTCAG 1800  
RESULT 14  
US-09-938-842A-1305  
; Sequence 1305, Application US/09938842A  
; Publication No. US20040009476A9  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Krepes, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SAME, AND METHODS OF USE  
; FILE REFERENCE: SCRIP1300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24



437 TCTAGAACCTACGCCGAGAGCTCTAGTTTCAACACACACTGATTCTTACTCCATGATG 496  
281 GLYAlaAlaAlaLysGlyGlyGlyGlyAlaAlaGlyAsp----- 293  
497 GCTGCTGGTGGCAGAACTCAAACTTTGGTCTCTGATGTTTATGGCCTTTTCTCAGCTTCA 556  
294 -----GluGlyGlyGlyAla----- 298  
557 AGAGGCCCAACTCAAGGCTCTTAAGTATGATGAAGATGGTGGGAAGCAAGTTTCAT 616  
299 ---CysGlyGlyGlyGlyGlyHisSerProGlnProGln----- 311  
617 TACCATGCTGCTGGTGAAGTGGCACTACCTCGCACCTAACCCCTGGCATGTTCTCTCC 676  
312 -----AlaValAlaValProAlaLysArg----- 319  
677 TCTAATGGGTCCAAAGTGTGTCTGCTTAATGCTTAATGCTCAAGAGGCTTAATGGCAGGCT 736  
320 -----LysAspLeuHisMetLeuValTrpSerSerAla 331  
737 CAGCTGAAGCTGAGGATGGATAGGACCTTCATATGTTGTTGGAGTTCAAGTGTCT 796  
332 SerProValSerGluArgAlaAlaValHisValPheGlyAlaGlyGlyAlaAlaAla 351  
797 TCACCAATTTCTGATGTTGTTGGTGGCCATGATGGA---GGAGGTCTATGATCAGAAA 853  
352 AspValLeuAlaLysGlyAlaGlnAlaTyrAspGluTyrGly-----ArgAsp 367  
854 GAAGTCAATGTAATGTAATCTCCAGAAAGTGGAGAATAATCATAGACACACTCAAGAA 913  
368 AspTyrSerSerArgThrLysAsnGlySerGlyGlyAlaAspLysGlyGlyProThrLeu 397  
914 GACTACCTAGAGAAAGATGATGTTGCTTGTGTAATAGAGAAATG----- 958  
388 SerLysLeuGlySerAsnSerThrAlaGlnLeuTyrProLysAspAspGlyGluGlyArg 407  
959 -----GACAGGGAGATGATCAGCTTGAAGGTGAGAGGTGGAGATGGGAAA 1006  
408 AlaAlaAlaValAlaMetProProAlaSerValMetThrArgLeuLeuLeuMetVal 427  
1007 CCAAAA-----ACCATGCTCCAGCAAGTGTGATGACAAGGCTTATATGATTATGCTG 1060  
428 TrpArgLysLeuLeuArgAsnProAsnThrTyrSerSerLeuLeuGlyValValTrpSer 447  
1061 TGGAGAAATCTCATCAAAACCCCAACACCTACTCTAGCCTTAATGCTCTCACTTGTCT 1120  
448 LeuValSerTyrArgTrpGlyLysGluMetProAlaAlaAlaAlaAlaAlaAlaAla 467  
1121 CTGTTTTCATCAAGTGAATGTGAGATGCTGCCATATAGCAAAAGTCTATCTCCATA 1180  
468 LeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeuGlyLeuPheMetAlaLeuGln 487  
1181 TTGTCAGACGCGAGGCTTGGCATGGCCATGTTTCTGCTCTCTTCTCATGGCTTTTGCAT 1240  
488 ProArgLeuLeuAlaCysGlyAsnLysLeuAlaAlaAlaAlaAlaAlaAlaAlaAla 507  
1241 CCGAGGTCATGACATGTAAGTGAATTCACAGCAGCTTTTGGCATGGCTGTGAGATTCCT 1300  
508 AlaGlyProAlaValMetAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 527  
1301 ACAGTCCAGCTGTCTGAGGAGCTGCTTCCATGCTTGGAGTCTCAAGGTGTCTCCTA 1360  
528 HisLeuAlaAlaValGlnAlaAlaLeuProGlnGlyLeuValProPheValPheAlaLys 547  
1361 CAGCTGGCATGTTTTCAGGAGCTTCTCCCAAGAGATTTGCCATTTGCTTTGCTTAAG 1420  
548 GluTyrGlyValHisProAspPheLeuSerThrAla-----TyrGly----- 561  
1421 GAATATATGATGATCTGATATTTCTCAGCAGCTGTTATTTTGGGATGCTGATGCT 1480  
562 ---ProLeuThr 564

517 SerIleAlaValGlyLeuArgGlyValLeuLeuHisIleAlaIleValGln 533  
1750 TCAATAGCAATGGTATTGAGGTGATCTCTCCATCATGCGCATGTTTCAG 1800

RESULT 15  
US-10-425-114-10077  
; Sequence 10077, Application US/10425114  
; Publication NO. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 10077  
; LENGTH: 1928  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700893164\_FLI  
US-10-425-114-10077

Alignment Scores:  
Pred. No.: 2,58e-122 Length: 1928  
Score: 1272.50 Matches: 289  
Percent Similarity: 63.74% Conservative: 45  
Best Local Similarity: 55.15% Mismatches: 101  
Query Match: 43.73% Indels: 89  
DB: 12 Gaps: 15

US-10-030-884-14 (1-573) x US-10-425-114-10077 (1-1928)

103 TrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThrLeuValMetGlyIlePro 122  
2 TGGGCGCATCAACCTGTGTTCTCTCTCCACCTCCCAACACCTTGTGATGGCATCCCT 61  
123 LeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeuMetValGlnValVal 142  
62 TTGCTCAAGGGATGATGGTGAATCTCTCA---GGGAGCCTCATGGTCAAAATTTGGTCT 118  
143 LeuGlnCysIleIleTrpTyrThrLeuMetLeuPheLeuPheGluTyrArgAlaAlaArg 162  
119 CTCAGTGCATCATTTGGTACACCTTGTCTCTCTTCTGAGTTTGAAGGTGCCAGA 178  
163 AlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaAlaAlaAlaAlaAlaAla 182  
179 ATGCTCATCTCTGAGCAGTTCCTGAC---ACTGCTGCTCCATGCTCTCCATCCATGTC 235  
183 AspSerAspValValSerLeu---AlaArgGlyAspValGluLeuGluAlaGluProAsp 201  
236 GACTCTGATGTCATGTCATTTGGATGGAAGAACACCACTTGAGACTGAAGCTGAGATCAAG 295  
202 GlyValAlaGlyAlaGlyAlaValSerSerArgGlyGlyAspAlaGlyArgValArgVal 221  
296 -----GAGATGTTAACTCCATGTC 316  
222 ThrValArgLysSerThrSerArgSerGluAlaAlaCysSerHisSerHisSer--- 240  
317 ACTGTGAGGAATCCCAATGCTTCAAGATCAGACATCTCTTCTAGAAGGTCTCAGGGTCTC 376  
241 GlnThrMetGlnProArgValSerAsnLeuSerGlyValGluIleTyrSerLeuGlnSer 260  
377 TCTTCCACCATCCAGCCCTTCCACCTTACCAATGCTGAGATATCTCTTTCATCC 436  
261 SerArgAsnProThrProArgGlySerSerPheAsnHisAlaAspPhePheAsnIleVal 280

Db 1481 TTGCCATAACT 1492

Search completed: March 3, 2004, 14:06:12  
Job time : 1368 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: March 3, 2004, 00:50:50 ; Search time 10669 Seconds  
(without alignments)  
11249.120 Million cell updates/sec  
Title: US-10-030-884-13  
Perfect score: 2769  
Sequence: 1 ccacgcgtccgtgcgcct.....aaaaaaaaaaaaaaaaaaag 2769

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 3470272 seqs, 21671516995 residues  
Total number of hits satisfying chosen parameters: 6940544  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_hg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_ma.\*  
20: em\_cm.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*  
29: em\_vi.\*  
30: em\_hg\_hum.\*  
31: em\_hg\_inv.\*  
32: em\_hg\_other.\*  
33: em\_hg\_mus.\*  
34: em\_hg\_pln.\*  
35: em\_hg\_rod.\*  
36: em\_hg\_mam.\*  
37: em\_hg\_vrt.\*  
38: em\_sv.\*  
39: em\_hgo\_hum.\*  
40: em\_hgo\_mus.\*  
41: em\_hgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2769	100.0	2769	6	AX046826	AX046826 Sequence
2	910.2	32.9	2233	8	AK103181	AK103181 Oryza sat
3	838.6	30.3	1926	6	AX653844	AX653844 Sequence
4	760	27.4	1686	6	AX652923	AX652923 Sequence
5	716.6	25.9	2126	8	AK063976	AK063976 Oryza sat
6	711	25.7	2470	8	AK103208	AK103208 Oryza sat
7	683	24.7	1884	6	AX653686	AX653686 Sequence
8	671.2	24.2	2358	8	AF056027	AF056027 Oryza sat
9	671.2	24.2	2402	8	AK102343	AK102343 Oryza sat
10	669.6	24.2	2343	8	AK099634	AK099634 Oryza sat
11	632.2	22.8	2457	8	AK101504	AK101504 Oryza sat
12	631.2	22.8	2276	8	AK101191	AK101191 Oryza sat
13	626.8	22.6	2293	6	AX046850	AX046850 Sequence
14	626.8	22.6	2293	8	BT008949	BT008949 Triticum
15	618	22.3	1827	6	AX654121	AX654121 Sequence
16	563	20.3	1845	6	AX652849	AX652849 Sequence
17	546.4	19.7	141025	2	AC123528	AC123528 Oryza sat
18	536.2	19.4	2162	6	AX046832	AX046832 Sequence
19	534.4	19.3	91883	8	CNS08CDY	EX000508 Oryza sat
20	484.8	17.5	1255	8	AK059577	AK059577 Oryza sat
21	470	17.0	170493	2	AP005769	AP005769 Oryza sat
22	452.6	16.3	504	6	AX046824	AX046824 Sequence
23	437.8	15.8	177049	8	AP003725	AP003725 Oryza sat
24	428.4	15.5	140844	2	AP005522	AP005522 Oryza sat
25	425.6	15.4	1426	6	AX046822	AX046822 Sequence
26	409.2	14.8	91870	8	AP004800	AP004800 Oryza sat
27	401	14.5	2297	8	AF515435	AF515435 Populus t
28	394.6	14.3	127130	2	AC098571	AC098571 Oryza sat
29	383	13.8	2316	8	BJU249298	AJ249298 Brassica
30	361.4	13.1	2264	8	AF087016	AF087016 Arabidops
31	352.2	12.7	2358	8	BT008691	BT008691 Arabidops
32	352.2	12.7	2404	8	AY091009	AJ249297 Brassica
33	351	12.7	2462	8	BJU249297	AJ132363 Brassica
34	348.6	12.6	2581	8	BJU132363	AY093960 Arabidops
35	331.4	12.0	1869	8	AF093960	AF372950 Arabidops
36	331.4	11.9	2261	8	AF372950	AF089084 Arabidops
37	329.8	11.9	2276	8	AF089084	A68597 Sequence 11
38	329.8	11.9	2292	6	A68597	A79382 Sequence 11
39	329.8	11.9	2292	6	A79382	AX046816 Sequence
40	323.8	11.7	1088	6	AX046816	AX506153 Sequence
41	319	11.5	1860	6	AX506153	AX651788 Sequence
42	319	11.5	2071	8	AF087820	AF087820 Arabidops
43	319	11.5	2071	8	AF087820	AX046842 Sequence
44	316.6	11.4	2549	6	AX046842	AF087819 Arabidops
45	305.6	11.0	1796	8	AF087819	

ALIGNMENTS

RESULT 1  
AX046826  
LOCUS  
DEFINITION Sequence 13 from Patent WO0068369.  
ACCESSION AX046826  
VERSION AX046826.1 GI:11876320  
KEYWORDS  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE  
AUTHORS Orozco, E.M., Weng, Z., Bruce, W.B., Cahoon, R.E. and Tao, Y.  
TITLE Auxin transport proteins

linear PAT 15-DEC-2000



```

QY 1981 AAAGAACCACTAATAAGCACTGATCTAGACAGCATCAGCATGATGATGCTGTTTTC 2040
Db 1981 AAAGAACCACTAATAAGCACTGATCTAGACAGCATCAGCATGATGATGCTGTTTTC 2040
QY 2041 TCCTGACAATCTGCATCTTTCTTACTACACAGTGTGCTTCTATCCATCCAGATGATCA 2100
Db 2041 TCCTGACAATCTGCATCTTTCTTACTACACAGTGTGCTTCTATCCATCCAGATGATCA 2100
QY 2101 TACACACATCTAGTGCATCTTTTCTTGTGATCTGCTGAGCTGATCTTCGGATGCT 2160
Db 2101 TACACACATCTAGTGCATCTTTTCTTGTGATCTGCTGAGCTGATCTTCGGATGCT 2160
QY 2161 CATCGCTCTGCCCATCACCCCTGCTTACTATACATCTTGTGCGGCTGTGAGCTCTCTCGC 2220
Db 2161 CATCGCTCTGCCCATCACCCCTGCTTACTATACATCTTGTGCGGCTGTGAGCTCTCTCGC 2220
QY 2221 TCGCTTCTTCAGCGTGCAGAGGCCCATGCTGTGTGTGTATCGGCCCCACATGAATTTTC 2280
Db 2221 TCGCTTCTTCAGCGTGCAGAGGCCCATGCTGTGTGTGTATCGGCCCCACATGAATTTTC 2280
QY 2281 TGATGAGATTTAGGATTTACTATTAGCTTAGCGAAGATGATGATGGTCTCGGCCCTGT 2340
Db 2281 TGATGAGATTTAGGATTTACTATTAGCTTAGCGAAGATGATGATGGTCTCGGCCCTGT 2340
QY 2341 CGGACTGCGGGAGTGCAGACAGACCCCTCGAACAAGTTTCTTTTGGCTTCTGTGCTTC 2400
Db 2341 CGGACTGCGGGAGTGCAGACAGACCCCTCGAACAAGTTTCTTTTGGCTTCTGTGCTTC 2400
QY 2401 GTCAGAAACAAAGTTTGGCTTTTGGATCGGCTCGAAGCTCGAAGCTCGAAGCTCGAAG 2460
Db 2401 GTCAGAAACAAAGTTTGGCTTTTGGATCGGCTCGAAGCTCGAAGCTCGAAGCTCGAAG 2460
QY 2461 ATCATCCATGAGATGATCTCTCTCGAATCTAGAGTACGAGGCAATTAAGATA 2520
Db 2461 ATCATCCATGAGATGATCTCTCTCGAATCTAGAGTACGAGGCAATTAAGATA 2520
QY 2521 CCACAGGCAATGGAATCAACAAAGCTTCATGCGAGCGCTATCATATCAAGGACACA 2580
Db 2521 CCACAGGCAATGGAATCAACAAAGCTTCATGCGAGCGCTATCATATCAAGGACACA 2580
QY 2581 TGCAGAAATCAACGGAGTCTAGTGGCAATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2640
Db 2581 TGCAGAAATCAACGGAGTCTAGTGGCAATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2640
QY 2641 GGGTTTCTAGACTGATTAAGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2700
Db 2641 GGGTTTCTAGACTGATTAAGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2700
QY 2701 ACAAATTTCTGCTTTTCTAGAAAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2760
Db 2701 ACAAATTTCTGCTTTTCTAGAAAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2760
QY 2761 AAAAAAAG 2769
Db 2761 AAAAAAAG 2769

RESULT 2
AK103181
LOCUS
DEFINITION
Oriza sativa (japonica cultivar-group) cDNA clone:J03121124, full
insert sequence.
ACCESSION
AK103181
VERSION
AK103181.1 GI:32988390
KEYWORDS
FLI_CDNA; CAP trapper.
SOURCE
Oriza sativa (japonica cultivar-group)
ORGANISM
Oriza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
1
The Rice Full-length cDNA Consortium, National Institute of
Agrobiological Sciences Rice Full-length cDNA Project Team;

```

```

Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,
Kojima, K., Namiaki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,
Ohtsuki, K., Shishiki, T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group, Ohtsuki, Y., Murakami, K.,
Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,
Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,
Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,
Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,
Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN,
Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,
Hara, A., Hashizume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,
Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y.,
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Yoshino, M., and Hayashizaki, Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice
Science 301 (5631), 376-379 (2003)
DOI: 10.1126/science.10869764
2 (bases 1 to 2233)
Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,
Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,
Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
Kanegawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kishimoto, N., Kobayashi, M.,
Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,
Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,
Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,
Namiaki, T., Narikawa, R., Niikura, J., Nishii, K., Nomura, K.,
Nomura, K., Ohtsuki, K., Ohtsuki, K., Ooka, M., Ooka, H.,
Osato, N., Ota, Y., Ohtsuki, K., Ohtsuki, K., Ohtsuki, K., Ohtsuki, K.,
Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,
Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,
Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Takami-Takeda, Y.,
Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,
Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,
Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and
Yoshimura, A.
Direct Submission
Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp,
Tel:81-29-838-7007, Fax:81-29-838-7007)
This clone is one of the 28K full-length cDNA clones from japonica
rice.
URL : http://cdna01.dna.affrc.go.jp/cdna/
NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K.,
Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,
Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiaki, T.,
Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and
Yamamoto, M.
FAIS Genome Sequencing & Analysis Group: Ohtsuki, Y., Iida, Y.,
Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,
Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,
Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S.,
Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,
Yoshimura, A., Matsubara, K. and Murakami, K.
Genome Exploration Research Group in Riken Genomic Sciences Center
and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,
Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,
Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,
Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,
Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishii, K., Nomura, K., Numasaki, R., Ohtsuki, N.,
Ohta, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,

```

Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,  
Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,  
Yasunishi, A. and Hayashizaki, Y.  
Location/Qualifiers

# FEATURES

source  
1. .2233  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="J033121124"

## ORIGIN

Query Match 32.9%; Score 910.2; DB 8; Length 2233;  
Best Local Similarity 77.3%; Pred. No. 2.7e-93;  
Matches 1318; Conservative 0; Mismatches 273; Indels 114; Gaps 13;  
QY 160 CATCCGGCCATGATCACCGCTGGACCTCTACCACTGTGACGGGGTGTGCCGC 219  
DB 211 CATCGCAGCATGATCACGGTGGTGGACCTGTACCACTGTGACGGGGTGTGCCGC 270  
QY 220 TGTACGTGGCATGACGCTGAGCTGCTCGTTCGGTGGTGGGCACTTTTCAGCCGG 279  
DB 271 TGTACGTGGCATGACGCTGAGCTGCTCGTTCGGTGGTGGGCACTTTTCAGCCGG 330  
QY 280 ACCAGTGTCTCGGGATCAACCGCTTCTGTCGGCTTCTGCGCTGTCCTCTCTCTCC 339  
DB 331 ACCAGTGTCTCGGGATCAACCGCTTCTGTCGGCTTCTGCGCTGTCCTCTCTCTCC 390  
QY 340 ACTTCACTCTCACCAACGACCTTTCGCCATGAACCTGCGCTTCTTGGCGCGGACAGCG 399  
DB 391 ACTTCACTCTCACCAACGACCTTTCGCCATGAACCTTCTGCGCGCGGACAGCG 450  
QY 400 TGCAAGAGTGGCGTCTGCGCTGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 459  
DB 451 TCAGAGAGTGCATGCTGCTGCGCTTCTGCGCTTCTGCGCTTCTGCGCTTCTGCGCT 502  
QY 460 GCGCGTCTGCGCTGCTGCGCTGCTGCGCTGCTGCGCTGCTGCGCTGCTGCGCTGCG 519  
DB 503 GCGCG-GGTCACTGCTGCTGCGCTGCTGCGCTGCTGCGCTGCTGCGCTGCTGCGCT 561  
QY 520 TCGTATGGCATCCGCTGCTGCGCTGCTGCGCTGCTGCGCTGCTGCGCTGCTGCGCT 570  
DB 562 TCGTATGGCATCCGCTGCTGCGCTGCTGCGCTGCTGCGCTGCTGCGCTGCTGCGCT 621  
QY 571 GCAGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 630  
DB 622 GCAGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 681  
QY 631 TCCTCTTGAAGTACCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 690  
DB 682 TCCTCTTGAAGTACCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 738  
QY 691 CCGGTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 750  
DB 739 CCGGTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 798  
QY 751 TCAGAGTTCAGAGCGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 810  
DB 799 GTG-----GAGCGCGGAGCTGCGAGCGGAGCGGAG-----GTGG 834  
QY 811 GGAAGCGCGCGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 870  
DB 835 GGAAGCGCGGAGTACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 894  
QY 871 CGTGTCTGCAAC-----TCGACTCCAGACCATGAGCGCGCGCTGCTGCTGCTGCTGCT 921  
DB 895 CGTGTCTGCAACCGGAGCGAGTACACTGCGAGTCCATGCGCGCGCGCTGCTGCTGCTGCT 954  
QY 922 CCGCGTGGAGATTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 981  
DB 955 CCGCGTGGAGATTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1014  
QY 982 TCAACCAAGCGCGCTTCTTCAACATGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1041

DB 1015 TCACCAAGCGCGAGTCTTCAACATCGTCGGCAAC-----GGCAAGC 1056  
QY 1042 CGGGGAGCGAGAGAGAGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1101  
DB 1057 AGGCGAGCGAGAGAGAGCGCG-----CGCGCGTGGCGGGGCGCACTGCCCGAGCG- 1109  
QY 1102 AGGCGTGGCG 1161  
DB 1110 -----GTGGTGGGAG 1158  
QY 1162 CCTCGCGCGTGTGCGAGCG 1221  
DB 1159 CCTCGCGCGTGTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG- 1215  
QY 1222 CCGAGCTCTCGCCAAAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1281  
DB 1216 GCGTGTGGCG 1275  
QY 1282 GCAGAGCAAG 1341  
DB 1276 TCGGGAACAAGAAC-----GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320  
QY 1342 GGTCCAACTCGAGCG 1401  
DB 1321 GGTCCAACTCGAGCG 1375  
QY 1402 TGGCGATGCG 1461  
DB 1376 -GGCGATGCG 1434  
QY 1462 TGATCCGGAGACCCCAACACCTACTCCAGCGCTCATCGCGCGCTGCTGCTGCTGCTGCT 1521  
DB 1435 TGATCAGAACCCCAACACCTACTCCAGCGCTCATCGCGCGCTGCTGCTGCTGCTGCT 1494  
QY 1522 ACAGTGGCG 1581  
DB 1495 ACAGTGGCG 1554  
QY 1582 CCGGTCTCGGAGTGGCG 1641  
DB 1555 CAGGCGTGGAGTGGCG 1614  
QY 1642 TCGCGTGGCGGAAACAAGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1701  
DB 1615 TTGCTGTGGGAACTCCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1674  
QY 1702 CCGTATGGCG 1761  
DB 1675 CCGTATGGCG 1734  
QY 1762 TCGTCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1821  
DB 1735 TTGTTTCCAGCGCGCTCTTCTTCCAGAGATCGTGCGCTTGTGTTTGTGTTTGTGTT 1794  
QY 1822 TTGTTTCCAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1846  
DB 1795 TTGTTTCCAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1819

## RESULT 3

AX653844  
LOCUS AX653844  
DEFINITION Sequence 3714 from Patent WO03000898.  
ACCESSION AX653844  
VERSION AX653844.1 GI:29156658  
KEYWORDS  
SOURCE Oryza sativa  
ORGANISM Oryza sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzeae; Oryza.  
REFERENCE 1



**FEATURES**

source Syngenta Participations AG (CH)  
Location/Qualifiers  
1..1686  
/organism="Oryza sativa"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:4530"

**ORIGIN**

```

Query Match      27.4%; Score 760; DB 6; Length 1686;
Best Local Similarity 74.4%; Pred.No.2.le-76;
Matches 1198; Conservative 1; Mismatches 256; Indels 155; Gaps 13;

QY 171 ATGATCACCGCGCTGGACCTTACCAGTGCTGACGGCGGTGTCGCCGTGTACGTGTTACGTGGCC 230
DB 1 ATGATCACGGTGTGGACTGTACCACTCTGACCGCGGTGTGCGTTGTACGTGGCG 60

QY 231 ATGACGTGTGGTACGGCTCCGTCCGTGTGTGGGCATCTTCACGCCGACCAAGTGCTCC 290
DB 61 ATGACGTGTGGTACGGCTCCGTCCGTGTGTGGGCATCTTCCTCCGCGACCAAGTGCTCC 120

QY 291 GGGATCAACCGCTTCGTGGCGCTCTTCGCGTGTCCGCTCCTCTCTCTTCCTTCACCTTCATCTCC 350
DB 121 GGCATCAACCGCTTCGTGGCGCTCTTCGCGTGTCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180

QY 351 ACCAACGACCCCCTTCGCGCATGAACCTTCGCTTCTTGCCGCCGACACGCTGCAGAAGTGG 410
DB 181 ACACAACCCCCTTCGCGCATGAACCTTCGCTTCTTCGCCGCCGACACGCTCCAGAAGTCC 240

QY 411 GCCTCTCTCGGCTGTGGCGCTGGCGCTCCGCGGCCTCTCTCTCCCGCGCGCGCTGGGG 470
DB 241 ATCTCTCTCGCCCTCTCTCGGCT-----CTGGTGGCGGCTCTCTCCGCGCGGGTCC--- 291

QY 471 CTCGACTGGAGCATCAGCTCTTCTCCTCTCC--ACGCTCCCCAAACGCTCGTCTATGGG 529
DB 292 CTCGACTGGCTCATCACCTCTTCTCCTCTCCAACTTCCCCAAMACCTCTGTCATGGG 351

QY 530 CATCCCGTGTCTGAGGCATGTACGGCGGTGTCTGGCGCGGACGCTCATGTGTCAGGT 589
DB 352 CATCCCGTGTCTAAGGGGAT----- 372

QY 590 CGTGTCTCCAGTGCACTACTCGTACAGCTCATGCTCTTCTCTTTCGAGTACCGCGC 649
DB 373 -----TGCATCATCTGGTACAGCTCATGCTGTCTCTTTCGAGTACCGTGG 419

650 CCGCGCGCGTCTCTCTCGACCAAGTTCCCGACGGCGCCGCGCTCCATGCTCTCTT 709
DB 420 CGCGCGGCTTCTTGTCTATGGAGAGTTCCCGACA---CGCGCGCTCCATCGTCTCTCTT 476

QY 710 CCGGTGCACTCCGACGTGCTCTCGTCCGACAGGGGGAAGTTCAGTCTCGAGCCCGAGCC 769
DB 477 CCGGTGGATTCCGACGTCGCTCTCGTCCGCGCGCGGTGTGGAAGCCG----- 526

QY 770 CGACGGCTCTCCGCGCGCGCGCGCTCTCTCTCCCGCGGGGACGCCCGCGGGTGG 829
DB 527 ---CGGAGCTCGAGCGGAGCGGAG-----GTCCGGGACGATGGCAAGATGCG 572

QY 830 CGTCACTCGGCGAAGTCCAACAGTTCGCGTCCGAGCGCGGTCTCGAC----- 881
DB 573 AGTCACCGTGGCGAAGTTCGACGAGTTCGCGTCCGAGCGCGGTCTCGACGCGGACGCA 632

QY 882 -TCGCACCTCCAGACCAATGACAGCCCCGTGTGTCAACTCTCCGGCGTGGAGATCTACTC 940
DB 633 GTCACTCTGAGTCCATGACGCGCGGTCTTCAAACCTCTCCGGCGTTCGAGATTIACTC 692

QY 941 GCTCAGTCTGTGCGCAACCAACCCCGCGGGTTCAGTTTCAACCAACGCGCGCTTCTT 1000
DB 693 GCTGCACTGTGCGGAACCGGACGCGCGGGGTTCAGCTTCAACCAACGCGCGGTTCTT 752

QY 1001 CAACATCTCTGGCGCGCGCCCAAGGAGGCGGAGGACGCGCGGGGACACGAGGAAAGGG 1060
DB 753 CAACATCTGTGGCAAC-----GGCAAGCACGCGCGACGAGGAAAGGG 794

QY 1061 CGCATCTCGCGCGCGCGGCGGAGGACACTCGCCGACCGCGAGCGCGTCTGCGCTGCCGCG 1120

```

Db	795	CGC-----CGCCGTTGGCGGGGCACCTCGCCGACCG-----GTGGTGGG	836
Qy	1121	CAAGAGAAAGAACTGTGCACATGTCTGTTGGAGCTCCAGCGCTCGCCCGTGTCCGAGCG	1180
Db	837	GAAGAGAAAGAACTGTGCACATGTCTGTTGGAGCTCAAGCGCTCGCCCGTGTGGAGCG	896
Qy	1181	CGCGCCGTGCAGCTCTTCGGCGCGCGCGCTGACCATGCGGAGCTCTCCCAAGG	1240
Db	897	CGCCGCGCGCGCGCGCTGCCTGCGCGGTGCACGTCTTCGGCGGTGTGGCGCCGACCA	956
Qy	1241	AGCCAGGCTTACGACGAGTACGGGCGCGACGACTACAGCAGCAGGACGAGAACGGGAG	1300
Db	957	CGGCGACGCCAAAGTGCTCAGCGCTATGATGAGTACAGCTTCGGGAACAAGAA-----	1010
Qy	1301	CGGCGGCGCGACAAGGCGGCGCGACGCTGTTCGAACTGGGGTCCAACTCGACGGCGCA	1360
Db	1011	-----TGAGNAGAACGGGCGCGACACTGTTCGAGCTGGGGTCCAACTCGACGGCGCA	1061
Qy	1361	GCTGTATCCCAAGAACGACCGGAGGGGAGGGCGGCGGGTGGCGATTCGCCCGCGCGAG	1420
Db	1062	GCTCCGCGCCAAAGAACGACGCGGAGGGGAGGGCGGC-----AGCGATTCGCCCGCGCGAG	1115
Qy	1421	CGTGATCAGCGGGTCACTCTCATCTGTTGTGGAGGAAGCTGATCCGGAACCCCAACAC	1480
Db	1116	CGTGATCAGCAGGCTCATATGATCATCTGCTGGAGGAAGCTGATCAGGAACCCAAACAC	1175
Qy	1481	CTACTCAGCTCATCGGCTGCTGGTCCCTGGTCTCTACAGTGGGGCATCGAGAT	1540
Db	1176	TTACTCAGGCTCTCTCGGTGTCATCTGGTCCCTCGTCTCTCTACAGTGGGGAAATTGAGAT	1235
Qy	1541	GCCAGCGCATCATCGCCGGTTCGATTTTCGATCTCTGTCGAGCGGGTCTCGGGATGGCCAT	1600
Db	1236	GCCGGCTATCATCGCCCGCTCGATTTTCGATCTCTTCAGTTCGAGGGCTCGGAATGGCCAT	1295
Qy	1601	GTTTCAGCCTAGCGCTGTTTCATGCGCTGACGCGGAGGATCATCGCTGTCGGGAACAAGCT	1660
Db	1296	GTTTCAGCCTAGGATTGTTTCATGCGCATTTGCAGCCACCGGATCATTCGCTGTGGGAATCCCTC	1355
Qy	1661	GGCGGCCATCGCGATGGCGGTTCGGTTCGTCGAGGCGCCCGCGGTCAATGGCGCGCGCTC	1720
Db	1356	TGCTTCGTATGCCATGCGCTCAGGTTCTCTGCTGTGGTCTTCGCGGTCAATGACTGGCGCTC	1415
Qy	1721	CATCGCGTCCGCTTCGCGCGCGTCTCTCTCCACATCGCCATGTCGAGG	1770
Db	1416	CATCGCGTCCGACTTCGCGGGTCTTCGTCATTCGCAATTGCCATTGTCAGG	1465

RESULT 5	AK063976	AK063976	2126 bp	mRNA	linear	PLN 24-JUL-2003
LOCUS	AK063976					
DEFINITION	Oryza sativa (japonica cultivar-group) cDNA clone:001-124-C02, full insert sequence.					
ACCESSION	AK063976					
VERSION	AK063976.1 GI:32973994					
KEYWORDS	FLI CDNA; oligo-capping.					
SOURCE	Oryza sativa (japonica cultivar-group)					
ORGANISM	Oryza sativa (japonica cultivar-group)					
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.					

1. The Rice Full-Length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-Length cDNA Project Team., Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shihishi, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group, Otomo, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,







Db	1054	ATGCCAAGAGCTCCACATGTTCTGTGGAGCTCGAGCGCGTCTCCCGTCTCAGAAGTCA	1113
QY	1183	CGCGCTGCACGCTTTTCGGCGCGCGCGCGCTGACCATGCCGACGCTCTCGCCAAAGGAG	1242
Db	1114	GGGCGCTGCCTGTGTTCACTGGCGCGCGCGCGCGCTCTCGACGTCGGCGCCAGGAA	1173
QY	1243	CCCA----GGCTACGACGAGTACGGCGCGACGACTACAGCAGCAGGACGAGAACGGGA	1299
Db	1174	TCCACATGCTCATCCCGCGCGAGCTCCGCGAGAACACCGCTCAGCGAACAGACGAGG	1233
QY	1300	GCGCGCGCGCGGACAAAGGGCGCGCGCGAGCTGTCGAAGCTTGGGCTCCAACTCGACGGCGC	1359
Db	1234	AGTACGCGCGAGTGGCATTTGGGTGGCGCGCGCGGCGGAGAACATTCAGCTCGACGGCGG	1293
QY	1360	AGCTGTACCCCAAGGACGA-----CGGCGAGGGGAGGG	1392
Db	1294	AGTGCACCCGAAGTCTGTCGAGCTCGACCGAACCGCGCGCGCGCGCGCGCGCGG	1353
QY	1393	CGCGCGCGGTGGCGATGTCGCGCGCGAGCGTATGACGCGGCTCATCTCATCATGCTGT	1452
Db	1354	CGGGGAGTACCAATGCGCGCGGAGCGTGAACACGCTCATCTCATATGCTGT	1413
QY	1453	GGAGGAAGCTGATCCGGAAACCCCAACCTACTCCAGCCTCATCGCGCTGTCTGTCTCC	1512
Db	1414	GGCGCAAGCTCATCCGCAACCCCAACACTTACTCCAGCCTCTCGCGCTCGCGTCTCC	1473
QY	1513	TGCTCTCTACAGGTGGGCGATCGAGATGCCAGGATCATCGCGCGTCTCATCTCATCTC	1572
Db	1474	TCGTCCGCTTCCGCTGGCAGCTTCCATGCCAGCAATCTCGAGAGATCATCTCCATTC	1533
QY	1573	TGTTCGACGCGGCTCTCGGGATGCGCATGTTTCAGCCTTAGCCCTGTTCATGGCGCTCGAGC	1632
Db	1534	TCTTCGACGCGAGCCTGGGGATGCGCATGTTTAGCCTGGGATTTGTTTCATGGCGCTCGAGC	1593
QY	1633	CGAGGATCATCGCGTTCGGGGAACAGGTGGCGGCAATCGGATGGGCGTTCGGTTCGTCTG	1692
Db	1594	CCAGCATCATCGCGGTGTGGCAATCATAGCCGCGCTCTCTCCATGGCGCTTCGCTTCTCG	1653
QY	1693	CAGCGCCCGGGTTCATGGCGCGCGCTCCATTCGCGCTCGGTCTGCGCGGCGTCTCTCTCTC	1752
Db	1654	CGGCGCTCGCGTCTGCGCGCGCGCTCAATCGCATCGGATTCGAGCTCCGCGGAGCCTCTCTC	1713
QY	1753	ACATCGCCATTCGTCAGGCTGCTCTGCTCCAGGGATTCGTCGCTTCGTTTCGCCAAGG	1812
Db	1714	ACGTCGCCATTGTTACGCGGCTCTACCAAGGAGTGTGCGCTTTGTTTTCGAAAAG	1773
QY	1813	AGTACGGGTTTCATCCGACATCCTCGACACGCG	1847
Db	1774	AATACATGTCACCGCGGCTCTCGAGCACGCG	1808

RESULT 6

AKI03208

LOCUS

DEFINITION

Oryza sativa (japonica cultivar-group) cDNA clone:J033122123, full insert sequence.

AKI03208

AKI03208.1 GI:32988417

VERSION

KEYWORDS

FLU DNA; CAP trapper.

ORYZA sativa (japonica cultivar-group)

ORYZA sativa (japonica cultivar-group)

SOURCE

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

1 The Rice Full-length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-length cDNA Project Team; Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohmeda, E., Yanagi, W., Suzuki, K., Li, C., Ohtsuka, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group; Otsomo, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,

/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="J033122123"

ORIGIN

Query Match 25.7%; Score 711; DB 8; Length 2470;  
Best Local Similarity 67.5%; Pred. No. 6.3e-71;  
Matches 1193; Conservative 0; Mismatches 430; Indels 144; Gaps 8;

171 ATGATACCGCGCTGGACCTCTACACGTGCTGACGCGGTGTCGCGCTGTACGTGGCC 230  
139 ATGATACCGCGCGGAGCTTCTACACGTGATGACGCGGTGTCGCGCTGTACGTGGCG 198  
231 ATGACGCTGGCGTACGGCTCGCTCGCTGGTGGCGCATCTTCAGCGCGGACCGATGCTCC 290  
199 ATGATCTCTGGCGTACGGGTGCTGAAGTGGTGGCGCATCTTCAGCGCGGACCGATGCTCC 258  
291 GGGATCAACCGCTTCTGGCGCTCTTTCGCGCTGCGCTCTCTCTCTCTCTCTCTCTCTCTCT 350  
259 GGGATCAACCGCTTCTGGCGCTCTTTCGCGCTGCGCTCTCTCTCTCTCTCTCTCTCTCTCT 318  
351 ACCAAGACCGCTTCTGGCGCATGAACCTGGCGCTTCTGGCGCGCGGACCGATGCTGAGGTG 410  
319 ACCAAGACCGCTTCTGGCGCATGAACCTGGCGCTTCTGGCGCGCGGACCGCTGAGAGCTC 378  
411 GCGCTCTCTGGCGCTTCTGGCGCTTCTGGCGCGCGCTTCTCTCTCTCTCTCTCTCTCTCTCT 470  
379 ATCGTCTCTGGCGCTTCT 429  
471 CTGAGCTGGAGATCAGCGCTTCT 530  
430 CTCGAGTGGACATCACCGCTTCT 489  
531 ATCCGCTGCTCGGAGGATGTCGCGCGCTGTCGCGCGCGCGCTCTCTCTCTCTCTCTCTCTCT 590  
490 ATCCGCTGCTCGGAGGATGTCGCGCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 546  
591 GTCGCTCTCTCGAGTGCATCTGTGATACAGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 650  
547 GTGGTGTCTCGAGTGCATCTGTGATACAGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 606  
651 GCGCGCGCTGCTCTCGAGGATGTCGCGCGCTGTCGCGCGCGCGCTCTCTCTCTCTCTCTCTCT 710  
607 GCCAGGATCTCTATCACGAGGATGTCGCGCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 663  
711 GCGCTGCTCTCGAGTGCATCTGTGATACAGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 770  
664 GTCGCTGAGCGCGAGTGTCTGTCTGATCGAGCGGAGGATGATCGAGACGAGGCG 723  
771 GACGCGTGTGCGCGCGCGCGCT 830  
724 GA-----GGTGAAGGAGGAGCGGCAAGATACAC 750  
831 GTCACCGTGGCGCAAGTCCACAGCTCGCGCTCCGAGGCGCGGTGCTCGCACTC---GCAC 887  
751 GTCACCGTGGCGCGCTCTCAACCGTCTGCGCTCTCGAGCTCTACTCGCGCGCTCTCATGGGG 810  
888 TCCAGACCATGACGCGCGCTGTGTCTCAACCTCTCTCGCGGTGGAGATCTACTCGCTCGAG 947  
811 TTCTCCAGCACACCGCGCGCGGAGCACTCTCAACACCGCGGAGATCTACTCGCTCGAG 870  
948 TCGTCCGCAACCCACCGCGCGGCTCCAGCTTCAACACCGCGGCTCTCTCAACATC 1007  
871 TCGTCCGCAACCCACCGCGCGGCTCCAGCTTCAACACCGGCTCTCTCAACATC 930  
1008 GTCGCGCGCGCGCGCGGAGGAGCGCGCGGCGGAGCGCGCGGCGGAGGAGGCGCGCGCGCGCG 1063  
931 GTCGCGCGGAGTCCCACTTCTCGCGCGGAGCGGCTTCTCGGCTGCGCGCGCGCGCGCGCGCG 990  
1064 -----ATGCG 1079  
991 CCGCGCGCGCTCCAACTACGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1050

QY 1080 GGAGGACACTCGCCGAGCGCGCGCTGCGCGTCCGCGCAAG----- 1124  
DB 1051 GGGCAGTACCCGCGCGCCAAACCCGCGCCATGGCGCGCGCCCAAGCCCAAGAGCGCGCC 1110  
QY 1125 -----AGGAGGACCTGSCACATGCTCTGAGCTCCAGC 1160  
DB 1111 AACCGCAGGCCAAGCGGAGGAGCGCAAGACCTCCACATGTTCTGTGAGCTCCAGC 1170  
QY 1161 GCGTCCCGCTGTCGAGCGCGCGCGCTGCGCGCTTTCGCGCGCGCGCGCGCTGACCAT 1220  
DB 1171 GCGTCCCGCTGTCGAGCGCTTTCGCGCAACCGCGCGCGAGTACAACGACGCGCGCGCGCTC 1230  
QY 1221 GCGCAGCTCTCGCCAAAGAGCGCGCGCTACGACGAGTACGCGCGCGCGCGCTACACG 1280  
DB 1231 AAGGAGTCCGATGCGCGCTGCGCTCGCGCGCAAGCGGCGCTGGAGGAGCGAC 1290  
QY 1281 AGCAGACGAAGAAAGAGGAGCGCGCGCGCGCAAGAGGCGCGCGCGCGCTGTCGAGCTG 1340  
DB 1291 TTCCAG-----CTTCGGGAACAGAGCGCGTTCGCGAGAGGAGCGCGCGCGCGCGAG 1344  
QY 1341 GGGTCCAACTCCAGCGCGCGCTGTACCCCAAGAGCGAGCGCGCGAGGAGCGCGCGCG 1400  
DB 1345 AAGAGGTGGCGCGCGCTGCGGTGAGCATGCGCAAGCTGGCTTGACGCGCGCGCG 1404  
QY 1401 GTGGGATCCCGCGCGCGCGTGTGATGACGCGCTCATCTCTCATGCTGTGAGGAG 1460  
DB 1405 ACGGGATCCCGCGCGCGCGTGTGATGACGCGCTCATCTCTCATGCTGTGAGGAG 1464  
QY 1461 CTGATCCGGAACCCCAACACCTACTCCAGCGCTCATCGCGCTGCTGGTCCCTGCTCTCC 1520  
DB 1465 CTGATCCGGAACCCCAACACCTACTCCAGCGCTCATCGCGCTCATGCTGTGCTCTG 1524  
QY 1521 TACAGTGGGGCATCGAGATCCAGATCCAGCTATCATCGCGCGCTGATTTGATTCCTGCGAC 1580  
DB 1525 TTCCGCTGGAATTCGAGATCGCGCGATCATCTCTGAAATCATCTCATCTCTCTCCGAC 1584  
QY 1581 GCGGCTCTGGATGCGCGCTTTCAGCTTACGCTTACGCTGCTCATGCGCTGCGCGCGAGGATC 1640  
DB 1585 GCGGCTCTGGATGCGCGCTTTCAGCTTACGCTTACGCTGCTCATGCGCTGCGCGCGAGATC 1644  
QY 1641 ATCGCTGCGGGAACAAAGCTGCGCGCATATCGCGGATGCGCGCTTCTGCTGAGGCGCC 1700  
DB 1645 ATCGCTGCGGGAACAAAGCTGCGCGCATATCGCGGATGCGCGCTTCTGCTGAGGCGCC 1704  
QY 1701 GCGGTCATGCGCGCGCGCTTTCATGCGCGCTGCTGCTGCGCGCGCTTCTCTCCACATCGC 1760  
DB 1705 GCGGTCATGCGCGCGCGCTTTCATGCGCGCTGCTGCTGCGCGCGCTTCTCTCCACATCGC 1764  
QY 1761 ATCGTCCAGGCTGCTGCTGCTCAGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1820  
DB 1765 ATCGTCCAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1824  
QY 1821 GTTCATCCCGACATCTGAGCACAGC 1847  
DB 1825 GTGCACCCGAGATCTCAGCACGCG 1851

RESULT 7  
AX653686  
LOCUS  
DEFINITION  
Sequence 3556 from Patent WO0300898.  
ACCESSION  
AX653686  
VERSION  
AX653686.1 GI:29156500  
KEYWORDS  
Oryza sativa  
SOURCE  
ORGANISM  
Oryza sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Bhrhartoideae; Oryzeae; Oryza.  
REFERENCE  
AUTHORS  
Chang, H. S., Chen, W., Cooper, B., Glazebrook, J., Goff, S. A., Hou, Y. M.,  
Katagiri, F., Qian, S., Tao, Y., Whitman, S., Xie, Z., Zhu, T. and Zou, G.

AX653686 1884 bp DNA linear PAT 22-MAR-2003  
Sequence 3556 from Patent WO0300898.  
AX653686  
AX653686.1 GI:29156500  
Oryza sativa  
Oryza sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Bhrhartoideae; Oryzeae; Oryza.  
REFERENCE  
AUTHORS  
Chang, H. S., Chen, W., Cooper, B., Glazebrook, J., Goff, S. A., Hou, Y. M.,  
Katagiri, F., Qian, S., Tao, Y., Whitman, S., Xie, Z., Zhu, T. and Zou, G.

TITLE Plant genes involved in defense against pathogens

JOURNAL Patent: WO 0300898-A 3556 03-JAN-2003;

Symgenia Participations AG (CH)

FEATURES Location/Qualifiers

source

1..1884

/organism="Oryza sativa"

/mol\_type="unassigned DNA"

/db\_xref="taxon:4530"

ORIGIN

Query Match 24.7%; Score 683; DB 6; Length 1884;

Best Local Similarity 67.1%; Pred. No. 9.3e-68;

Matches 1159; Conservative 2; Mismatches 422; Indels 144; Gaps 8;

Qy 211 TGTGTCGGCTGTACGTGGGCGCATGACGCTGGCTGACGGCTCCGCTCGCTGGTGGCGCATCT 270

Db 2 TGTGTCGGCTGTACGTGGGCGCATGACGCTGGCTGACGGCTCGGTGAAGTGGTGGCGCATCT 61

Qy 271 TCACGCGGACACAGTCTCCGGGATCAACCGCTTGTGGCGCTCTTCGGCGTGGCGCTCC 330

Db 62 TCACGCGGACACAGTCTCCGGGATCAACCGCTTGTGGCGCTCTTCGGCGTGGCGCTCC 121

Qy 331 TCTCTCTTCCATCTTCACTCCACCAAGACCCCTTGGCCATGAACCTTGGCTTCTTGGCGG 390

Db 122 TCTCTCTTCCATCTTCACTCCACCAAGACCCCTTGGCCATGAACCTTGGCTTCTTGGCGG 181

Qy 391 CGACACGCTGTGAGAAAGTGGCGTCTCGCGTGTGGCGCTGGCGTCCCGCGGCTCT 450

Db 182 CGACACGCTGTGAGAAAGTGGCGTCTCGCGTGTGGCGCTGGCGTCCCGCGGCTCT 232

Qy 451 CTTCCCGCGCGGCTCGGCGTCTGAGATGATGATGATGATGATGATGATGATGATGATGAT 510

Db 233 ACCTTCCCGCGCGGCTCGGCGTCTGAGATGATGATGATGATGATGATGATGATGATGATGAT 292

Qy 511 CCAACACGCTGTATGGGATCCCGCTCTGCGAGGATGATGATGATGATGATGATGATGATGATGAT 570

Db 293 CCAACACGCTGTATGGGATCCCGCTCTGCGAGGATGATGATGATGATGATGATGATGATGATGAT 349

Qy 571 GCACGCTCATGTGTCAGTGTGTGCTCTGAGTGTGATGATGATGATGATGATGATGATGATGATGAT 630

Db 350 GTAGCTCATGTGTCAGTGTGTGCTCTGAGTGTGATGATGATGATGATGATGATGATGATGATGAT 409

Qy 631 TCCCTCTTCAGTATCCCGCGCGCGCGCTGTCTCTCGACCAAGTATCCCGCGCGCGCGG 690

Db 410 TCATGTTTCAGTATCCCGCGCGCGCGCTGTCTCTCGACCAAGTATCCCGCGCGCGCGG 469

Qy 691 CCGGCTCATGTCTCTTCGGGTCGATCTCGACGTGTCTCTCGTCTCGTCCGACGGGGAGG 750

Db 470 GCG---CCATCGGCTCATGTCTGTCGACGCGACGCTGTCTGTCGACGCGGGGAGGG 526

Qy 751 TCAGGCTTCGAGGCGGAGCGCGCGCGCGCTGTCTCTCGACCAAGTATCCCGCGCGCGG 810

Db 527 ACATGATTCGAGACGGAGGCGAGGTGNA----- 554

Qy 811 GGACGCGCGGCGGGTGGCGGTCAACGCTGGCGAAGTTCACAGCTCGCGCTCCGAGGCGG 870

Db 555 -GGAGGACGCGAAGATACAGTCAACGCTGGCGCGCTCCAAACGCTCGCGCTCCGACGCT 613

Qy 871 CGTGTCTGCATC---GCATCTCCACAGACCATGACGCGCGCTGTGTCCAACTCTCCGGG 927

Db 614 ACTCGCGCGCTCATGTTGGGTTCTCAGACACAGCGCGCGCGGCGAGCACTTCAACAG 673

Qy 928 TGGAGATCTACTCGTGTGAGTGTGCGGCAACCCACCGCGCGGCTCCAGTTCAACC 987

Db 674 CCGAGATCTACTCGTGTGAGTGTGCGGCAACCCACCGCGCGGCTCCAGTTCAACC 733

Qy 988 ACGCCGATCTTCAACATGTCGGCGCGCGCGCAAGGGAGCGGAGGCGGGGG 1047

Db 734 ACATGATCTTCACTCCATGTCGGGCGAGTCCAACTTCGCGCGCGGGAGCGGTTG 793

Qy 1048 ACGAGGAGAGGGGCG----- 1063

Db 794 GGGTGGCACCGGCGCGCGCGCGCTCCAACTACGAGGAGGAGCGGCGGCGGCCCA 853

Qy 1064 ----ATGCGCGCGCGCGCGGAGGACACTCGCGCGCAGCGCGCAGCGCTCGCGTCCCGG 1119

Db 854 ACAAGCGCGCAGCAAGTACGGCAGTACCCGCGGCCACCCGCGCAATGGCGCGCGC 913

Qy 1120 CCAAG-----AGGAGGACCTGCACA 1140

Db 914 CCAAGCCCAAGAGGCGGCCCAACGCGCAGGCCAAGGCGGAGGACGCAAGACCTCCACA 973

Qy 1141 TGTCTGCTGAGCTCCAGCGCTCGCGCTGTCGAGCGCGCGCGCTGTCACGCTCTTCG 1200

Db 974 TGTCTGCTGAGCTCCAGCGCTCGCGCTGTCGAGCGCTGTCGAGCGCTTTCGCAACGGCGCGAGT 1033

Qy 1201 GCGCGCGCGCGCTGACCATCCGACGCTCTCCCAAGAGAGCCAGGCTTACGACGAGT 1260

Db 1034 ACAACGACGCGCGCGCTCAAGGAGTCCGATGCGCGCTCGCGCGCAAGCGG 1093

Qy 1261 ACGGCGCGGACTACAGCAGCAGGAGCAAGAAACGGGAGCGCGCGCGCAAGGGCG 1320

Db 1094 ACGGCTGGAGAGGAGCAGCTTCACTTCCGGAACAGAGCGCTCGCGAGAGGAGCGCG 1153

Qy 1321 GCGCGACGCTGTCCAAAGTGGGTCCAACTCGACGCGCGAGCTGTATCCCAAGAGCAGC 1380

Db 1154 AGCGCGCGCAGAGAGAGCGTGGCGCGCGCTGTGGGTGAGCATGGCAAGCTTGGCT 1213

Qy 1381 GCGAGGAGGCGCGCGCGCTGCGCGCTGCGCGCGCGCGCGCGCGCTGATGACGCGCTCATCC 1440

Db 1214 T-----GACCGCGCGCGCGCGCGCTGCGCGCGCGCTGCGCGCGCTGATGACGCGCTCATCC 1267

Qy 1441 TCATCATGTGTGGAGGAAGCTGATCCGGAACCCCAACACCTACTCCAGCTCATCGCGG 1500

Db 1268 TGATCATGTGTGGCGCAGCTCATCCGCAACCCCAACACCTACTCCAGCTCATCGCGG 1327

Qy 1501 TCGTCTGCTCCCTGCTCTCTACAGTGGGGCATCGAGTCCGAGCGCGCGCGCATCGCGCGG 1560

Db 1328 TCATCTGCTCGCTGCTCTCTCGGTGGAATTCGAGATGCGCGCGCGCATCTCTGAAT 1387

Qy 1561 CGATTTCCATCTCTCGACGCGCGCTCTCGCGGATGCGCATGTTTCAGCTAGGCTGTCTCA 1620

Db 1388 CCATCTCCATCTCTCGACGCGCGCTCGCATGTTTCAGCTCGGCTGTCTCA 1447

Qy 1621 TGGCGCTGCGCGGAGGATCATCGTGGCGGAACAGCTGGCGGCGCATCGCGAGCGGCG 1680

Db 1448 TGGCGCTGCGCGGCGGATCATCGTGGCGGAACAGCTGGCGGAGCTTCGCGCATGCGCG 1507

Qy 1681 TCGGTTTCGTGCGAGCGCGCGCTCATGCGCGCGCTCCATCGCGCTCGGTCTGCGCG 1740

Db 1508 TCGGTTTCGTGCGCGCGCGCGCTCATGCGCGCGCTCCATCGCGCTCGGTCTGCGCG 1567

Qy 1741 GCGTCTCTTCCACATCGCCATCGTCCAGGCTGTCTGCTTCAGGAGGATCGTGGCGTTCG 1800

Db 1568 GCAGCTCTCTCCAGCTCGCCATCGTGCAGGACGCGTTGCCMCASGGCATTTGCCCTTSG 1627

Qy 1801 TGTTCGCGCAAGGAGTACGCGCTTCATCCCGACATCTCTGAGCACAGCG 1847

Db 1628 TCTTCGCGCAAGGAGTACGCGCTTCATCCCGACATCTCTGAGCACAGCG 1674

# RESULT 8

AF056027

LOCUS

DEFINITION

AF056027

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

AUTHORS

AF056027 2358 bp mRNA linear PLN 03-AUG-1998

Oryza sativa auxin transport protein REH1 (REH1) mRNA, complete cds.

AF056027.1 GI:3377508

Oryza sativa

Oryza sativa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

AUTHORS

1 (bases 1 to 2358)

Luehnig,C., Gaxiola,R.A., Grisafi,P. and Fink,G.R.





Db 496 ATGGGGATCCCTTTCTCAAGGGCATGTACGGGAGTTCT---CGGGAGCCTCATGGTG 552  
 Qy 585 CAGGTCTGTCCTCCAGTGCATCTCTGGTACAGCTCATGCTCTTCTTCTTCCAGTAC 644  
 Db 553 CAGATCTGTCGTCTGATGTCATATCTGGTACAGCTCATGCTCTTCTTCTTCCAGTAC 612  
 Qy 645 CGCGCCGCGCGCGCTCTCTCTCCAGCAGTTTCCCGACGGCGCGCGCTCTTCCAGTCTC 704  
 Db 613 CGCGCGCGCGCGATCTCATACCGAGCAGTTCCCGGA--CAGCGCGCCCAACATCGCC 669  
 Qy 705 TCCTTCCTGCTGATCTCGAAGTCTCTCTCTGCTGCGCAGGGGGAGCTGAGCTCAGAGCC 764  
 Db 670 TCCATCTGTCGTGACCCCGGAGTCTGTCTGTGAGACGCGAGGAGGAGCCATCAGAGC 729  
 Qy 765 GAGCCCGACGGGCTCGCGCGCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 824  
 Db 730 GAGACGGA-----GGTGAAGGAGGAGCGGAGG 756  
 Qy 825 GTGCGCGTCAACGTCGCGAAAGTCCACAGCTCGGCTCCGAGGCGCGGTCTCGCACTC- 883  
 Db 757 ATACACGTCACGTCGCGCGCTTCCAAACGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 816  
 Qy 884 --GCACCTCCAGACCATGACGCGCGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 941  
 Db 817 ATGGCTTTCTCAGCACACGCGCGCGCGCGGACCTTCAACAGCGCGAGATCTACTCG 876  
 Qy 942 CTGCACTGTCGCGCAACCCCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1001  
 Db 877 CTGCACTGTCGCGCAACCCCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 936  
 Qy 1002 AACATCTGTCG 1061  
 Db 937 TCCATGTCGTCG 996  
 Qy 1062 GC----- 1063  
 Db 997 GCACGCGCGCGCGCTTCCACTACGAGGAGCGGCTCCAGCCCAAGTACCGGCTCCG 1056  
 Qy 1064 -----ATGCGCGCGCGCGCGCGGAGAC-ATGCGCGCGCGCGCGCGCGCGCGCG 1111  
 Db 1057 GGTGCAATGTCG 1116  
 Qy 1112 -----CGTGGCG 1139  
 Db 1117 GGGCCCAAGGGCGCGCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1176  
 Qy 1140 ATGCTCTGTCGAGCTCCAGCGCGCTCGCGCGCTGTCGAGCGCGCGCGCGCGCGCG 1199  
 Db 1177 ATGTTCTGTCGAGCTCCAGCGCGCTCGCGCGCTGTCGCGCGCGCGCGCGCGCGCG 1221  
 Qy 1200 GCGCGCGCGCGCGCTGACCATGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1259  
 Db 1222 GCG 1281  
 Qy 1260 TACG 1319  
 Db 1282 GGAGCGAAGGACAGGAGGAGTACGTCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1341  
 Qy 1320 G---GGCCGACGCTGTCGAGCTGGGTCCTCACTCGAGCGCGCGCGCGCGCGCGCG 1376  
 Db 1342 GTCATGGAAGGAGCG 1401  
 Qy 1377 GACGCGCGAGGAGCG 1436  
 Db 1402 CCAGCAAGGCG 1461  
 Qy 1437 ATCTCTCATGTCGTCGAGGAGTGTATCGGAGACCGCAACCTCTCTCTCTCTCTCTCT 1496  
 Db 1462 ATCTCTCATGTCGTCGAGGAGTGTATCGGAGACCGCAACCTCTCTCTCTCTCTCTCT 1521  
 Qy 1497 GCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1556

Db 1522 GGCTCATCTGTCCTCTGCTTTCAGTGGAACTTCGAGATGCGCGCCATCGTCTCTG 1581  
 Qy 1557 CGGTTCATTCGATCTCTGTCGAGCGCGGTCTCTCGGATGCGCCATGTTTACGCTTAGCGCTG 1616  
 Db 1582 AATTCATTCGATCTCTGTCGAGCGCGGTCTCTCGGATGCGCCATGTTTACGCTTAGCGCTG 1641  
 Qy 1617 TTTATGCGCTCTGTCGAGCGCGGTCTCTGTCGAGCGCGGTCTCTCGGATGCGCCATCGCGATG 1676  
 Db 1642 TTTATGCGCTCTGTCGAGCGCGGTCTCTGTCGAGCGCGGTCTCTCGGATGCGCCATCGCGATG 1701  
 Qy 1677 GCGTCTGCTCTGTCGAGCGCGGTCTCTGTCGAGCGCGGTCTCTCGGATGCGCCATCGCGATG 1736  
 Db 1702 GCGTCTGCTCTGTCGAGCGCGGTCTCTGTCGAGCGCGGTCTCTCGGATGCGCCATCGCGATG 1761  
 Qy 1737 GCGTCTGCTCTGTCGAGCGCGGTCTCTGTCGAGCGCGGTCTCTCGGATGCGCCATCGCGATG 1796  
 Db 1762 GCGTCTGCTCTGTCGAGCGCGGTCTCTGTCGAGCGCGGTCTCTCGGATGCGCCATCGCGATG 1821  
 Qy 1797 TTTGCTCTGTCGAGCGCGGTCTCTGTCGAGCGCGGTCTCTCGGATGCGCCATCGCGATG 1846  
 Db 1822 TTTGCTCTGTCGAGCGCGGTCTCTGTCGAGCGCGGTCTCTCGGATGCGCCATCGCGATG 1871

RESULT 10  
 AK099634 2343 bp mRNA linear PLN 24-JUL-2003  
 LOCUS Oryza sativa (japonica cultivar-group) cDNA clone:J013059F15, full  
 DEFINITION Insert sequence.  
 ACCESSION AK099634  
 VERSION AK099634.1 GI:32984843  
 KEYWORDS FLJ CDNA; CAP trapper.  
 SOURCE Oryza sativa (japonica cultivar-group)  
 ORGANISM Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzeae; Oryza.

REFERENCE 1  
 AUTHORS The Rice Full-length cDNA Consortium, National Institute of  
 Agrobiological Sciences Rice Full-length cDNA Project Team;  
 Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,  
 Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,  
 Kojima,K., Nami,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,  
 Ohtsuki,K., Shishiki,T., Foundation of Advancement of International  
 Science Genome Sequencing & Analysis Group; Ohtsuki,Y., Murakami,K.,  
 Tida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,  
 Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,C., Lu,M.,  
 Nariawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J.,  
 Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,  
 Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN;  
 Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,  
 Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,  
 Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y.,  
 Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,  
 Yoshino,M. and Hayashizaki,Y.  
 TITLE Collection, mapping, and annotation of over 28,000 cDNA clones from  
 Japonica rice  
 JOURNAL Science 301 (5631), 376-379 (2003)  
 MEDLINE 22752273  
 PUBMED 12869764

REFERENCE 2 (bases 1 to 2343)  
 AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,  
 Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,  
 Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T.,  
 Hori,F., Hotta,I., Iida,J., Ikeda,R., Imamura,K., Imoto,M.,  
 Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,  
 Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M.,  
 Kodama,T., Kojima,K., Kondo,S., Konno,H., Kouda,M.,  
 Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M.,  
 Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A.,  
 Mizuno,K., Murakami,K., Murata,M., Negata,T., Nakamura,M.,  
 Namiaki,T., Nariawa,R., Niikura,J., Nishi,K., Nomura,K.,  
 Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H.,



Osato, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shingawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

Direct Submission

Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel: 81-29-838-7007, Fax: 81-29-838-7007)

This clone is one of the 28K full-length cDNA clones from japonica rice.

URL: <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.

## COMMENT

FAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y., Yoshimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Nariakawa, R., Niihara, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

Location/Qualifiers

1. 2343

/organism="Oryza sativa (japonica cultivar-group)"

/mol\_type="mRNA"

/cultivar="Nipponbare"

/db\_xref="taxon:39947"

/clone="J013059F15"

## FEATURES

source

## ORIGIN

Query Match 24.2%; Score 669.6; DB 8; Length 2343;  
Best Local Similarity 66.8%; Pred. No. 2.8e-66;  
Matches 1195; Conservative 0; Mismatches 424; Indels 171; Gaps 10;

Qy 165 GCGGCCATGATACCGCGTGGACCTCTACCGTCTGCTACCGCGTGGTGGCTGTAC 224

Db 145 GCGAAGATGATACCGCGCGGCGACTTCTACCGTATGACGCGGATGGTGGCTGTAC 204

Qy 225 GTGGCCATGACGCTGGCGTACCGCTCCGTCGGTGGTGGCGCATCTTACGCGGACCGAG 284

Db 205 GTGGCGATGATCTGGCGTACCGCTCCGTCAGTGGTGGCGCATCTTACGCGGACCGAG 264

Qy 285 TGCTCCGGGATCAACCGCTCGTGGCGCTTTCGCGTCCGCTCCCTCTCTCCATCTTC 344

Db 265 TGCTCCGGGATCAACCGCTCGTGGCGCTTTCGCGTCCGCTCCCTCTCTCTCCATCTTC 324

Qy 345 ATCTCCACCAACGACCCCTTCGCGCATGACCTGCGTTCCTGGCGCGCGACACGCTGACG 404

Db 325 ATCTCCACCAACGACCCGTTACGATGACCTCGGTTTCATCGCGCGCGACACGCTGACG 384

Qy 405 AAGTGGCCGCTCTCGCGCTGTGGCGCTGGCGCTCCCGCGGCTCTCTCTCCCGCGCGCG 464

Db 385 AAGTGATGTGTGTGGCCATGCTCACGGCGTGGAGCCACCTCAGC-----CGCGCG 435

Qy 465 CTCGGGCTCGACTGGAGCATCACGCTCTTCTCCCTCTTCCAGCTTCCCAACACGCTCGTC 524

Db 436 GGGAGCCTCGAGTGGAGCATCACGCTCTTCTCCCTCTTCCACGCTGCCCAACACGCTCGTC 495

Qy 525 ATGGGCATCCGCTGCTGCGGAGCATGTACGGCGCGTCTGCGCGCGCGCGCTCATCGTC 584

Db 496 ATGGGGATCCCTTTGCTCAAGGGCATGTACGGGGATTCT---CCGACGCTCATGGTG 552

Qy 585 CAGTCTCGTCTCTCCAGTGCATCATCTGTACACGCTCAAGCTCTTCTCTTCGAGTAC 644

Db 553 CAGATCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 612

Qy 645 GCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 704

Db 613 GCGCGCGCGCGCGATGCTCATCACGAGCAGTCTCCCGA---CACCAGCGCAACATCGCC 669

Qy 705 TCTTTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 764

Db 670 TCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 729

Qy 765 GAGCCCGACGGCGTGGCGCGCGCGCGCGCTCTCTCCCGCGCGCGCGCGCGCGCGCGCG 824

Db 730 GAGACGGA-----GGTGAAGGAGGAGCGGCGG 756

Qy 825 GTGGCGCTCACCGTGGCGCAAGTCCACGAGTCTGCGCTTCCGAGCGCGCGTCTGCTGCTGCTC- 883

Db 757 ATACAGCTCACCGTGGCGCGCTCCAAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 816

Qy 884 --GCATCTCCAGACCATGACGCGCGTGTCTCAACTCTTCCGCGTGGAGATCTACTCG 941

Db 817 ATGGGCTTCTCCAGCACACGCGCGCGCGCGAGCAACTTCAACACGCGGAGATCTACTCG 876

Qy 942 CTGCAGTCTGCGCGCAACCCACCGCGCGCGTCTCAAGCTTCAACACGCGGAGTCTTCTTC 1001

Db 877 CTGCAGTCTGCGCGCAACCCACGCGCGCGGTTCAAGCTTCAACACGCGGAGTCTTCTTC 936

Qy 1002 AACATCTGCGCGCGCGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1061

Db 937 TCCATGTTGGCGCGCGAGTCTCAACTTTCGCGCGCGCGCGCGCTTTCGCGCGTCCGACCGCG 996

Qy 1062 GC-----CGTCCG 1063

Db 997 GCGCGCGCGCGCGCTCAACTAGAGGAGCGCGCTTCAAGCGCAAGTACCGCTCCCG 1056

Qy 1064 -----ATGCGCGCGCGCGCGCGAGGAC-ATGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1111

Db 1057 GCGTGAATGCG 1116

Qy 1112 -----CGTCCG 1139

Db 1117 GCG 1176

Qy 1140 ATGCTCGTCTGGAGCTCCAGCGCTTCCCGCTGTCGAGCGCGCGCGCGCGCGCGCGCGCGCG 1199

Db 1177 ATGTTCTGCTGGAGCTCCAGCGCTGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1221

Qy 1200 GCGCGCGCGCGCGTGGACCATGCGCGCTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1259

Db 1222 GCG 1281

Qy 1260 TACGGCG 1319

Db 1282 GGAGCGAAGGAGCG 1341

Qy 1320 G---GGCGGAGCGTGTGGAAGCTGGGTCCAACTCGACGCGCGCGCGCGCGCGCGCGCGCG 1376

Db 1342 GTCATGAGCAGGAGCG 1401

Qy 1377 GACCGCGAGGAGCG 1436

Db 1402 CCGAGCAGGCG 1461



1437 ATCCCTCATGCTGGTGGAGGAGCTGATCCGGACCCACACCTTACTCCAGGCTCATC 1496  
 Db 1462 ATCCCTGATCATGGTGGGCGAGCTATCCGACACCACTTACTCCAGGCTCATC 1521  
 Qy 1497 GCGGTGCTGGTCCCTGGTCTTCTACAGTGGGGCATCGAGATCCAGGATCATCGCC 1556  
 Db 1522 GCGCTCATGCTGGTCCCTGGTCTTCTACAGTGGGGCATCGAGATCCAGGATCATCGCC 1581  
 Qy 1557 CGGTGATTTTCATCTCTCGGACGCGGCTCTCGGATGGCCATGTTTCAGCTAGGCTG 1616  
 Db 1582 AAATCCATCTGATCTCTCGGACGCGGCTCTCGGATGGCCATGTTTCAGCTAGGCTG 1641  
 Qy 1617 TTCATGGCGCTGCAGCCGAGGATCATCGCTGCGGGACACAGCTGGGGCCATCGGATG 1676  
 Db 1642 TTCATGGCGCTGCAGCCGAGGATCATCGCTGCGGGACACAGCTGGGGCCATCGGATG 1701  
 Qy 1677 GCGTCCGCTGTCGCGAGGCGCGGCTCATGCGCGGCTCCATCGCGTGGCTGCTG 1736  
 Db 1702 GCGTGGCGGCTTCTGCGGCGCGGCTCATGCGCGGCTCCATCGCGTGGCTGCTG 1761  
 Qy 1737 CGCGCGCTTCTCTCCATCGCTCCAGGCTGCTGCGCTCCAGGATCGTCCCG 1796  
 Db 1762 CGTGGCAGCTCTCTGACGCTGCGGATGCTGCGGAGCTGCTGCGGAGGATGTCGCC 1821  
 Qy 1797 TTCGTGTCGCGAGGATGAGGCTTCTATCCGACATCTCTGACACAGC 1846  
 Db 1822 TTCGTGTCGCGAGGATGAGGCTTCTATCCGACATCTCTGACACAGC 1871

RESULT 11  
 LOCUS AK101504 2457 bp mRNA linear PLN 24-JUN-2003  
 DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:J033044E23, full insert sequence.  
 ACCESSION AK101504  
 VERSION AK101504.1 GI:32986713  
 KEYWORDS FLI CDNA; CAP trapper.  
 SOURCE Oryza sativa (japonica cultivar-group)  
 ORGANISM Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Eriarthroideae; Oryzaceae; Oryza.

1 The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team; Kikuchi, T., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Kojima, N., Yamada, H., Ooka, H., Hotta, I., Ohseuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group; Ootomo, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Oeato, N., Oca, Y., Sato, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.

TITLE japonica rice  
 JOURNAL Science 301 (5631), 376-379 (2003)  
 MEDLINE 22752273  
 PUBMED 12869764  
 REFERENCE 2 (bases 1 to 2457)  
 AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashidume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoaka, T., Hori, F., Hotta, I., Iida, J., Iida, J., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katch, H., Kawagashira, N., Kawai, J., Kawamata, M.,

Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kuroaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K., Ooka, H., Oeato, N., Ota, Y., Ootomo, Y., Ooka, H., Ohtsuka, K., Sakai, K., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, P., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toyama, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yoshimura, A., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

TITLE Direct Submission  
 JOURNAL Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (e-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)  
 COMMENT This clone is one of the 28K full-length cDNA clones from japonica rice.  
 URL : http://cdna01.dna.affrc.go.jp/cDNA/  
 NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, B., Yahagi, W., Suzuki, K., Li, C., Ohtsuka, K., Shishiki, T. and Yamamoto, M.

PAIS Genome Sequencing & Analysis Group: Ootomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashidume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katch, H., Kawai, J., Kouda, M., Kishikawa-Hirozane, T., Kojima, K., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ooka, H., Oeato, N., Ota, Y., Satoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, P., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toyama, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

Location/Qualifiers  
 1..2457  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /clone="J033044E23"

FEATURES  
 source  
 Query Match 22.8%; Score 632.2; DB 8; Length 2457;  
 Best Local Similarity 63.2%; Pred. No. 4.5e-62;  
 Matches 1141; Conservative 0; Mismatches 538; Indels 126; Gaps 6;

ORIGIN  
 Qy 163 CCCTGGCCATGATCAGCGCTGGGCTACGGCTCCGCTGGTGGCGGCATCTTCACGCCGGACC 222  
 Db 344 CTGAGGAGATGATATCCGGGCACGACTTCTACACGGTGATCGCGGGTGGTGGCGCTGT 403  
 Qy 223 ACCTGGCCATGATCAGCGCTGGGCTACGGCTCCGCTGGTGGCGGCATCTTCACGCCGGACC 282  
 Db 404 ACCTGGCGATGTTCTGGGCTACGGGTGGTGGCGGCATCTTCACGCCGGACC 463  
 Qy 283 AGTGTCCGGGATCAACCGCTTCGTGGGCGCTCTTCCGCCGTGGCGCTCTCTCTCCACT 342  
 Db 464 AGTGTCCGGGATCAACCGCTTCGTGGGCGCTCTTCCGCCGTGGCGCTCTCTCTCCACT 523

343 TCATCTCCACCAACGACCCCTTTGGCAATGAACCTCGCTTCTTGGCGCGCCGACACGCTGC 402  
Db  
524 TCATCTCCACCAACGACCCCTTTGGCAATGAACCTCGCTTCTTGGCGCGCCGACACGCTGC 583  
Qy  
403 AGAAGGTGGCCCTCTCGGCTGCTGGGCTCGGCTTCCCGGGGCTCTCTCTCCCGGGCG 462  
Db  
584 AGAAGGTGGCCCTCTCGGCTGCTGGGCTCGGCTTCCCGGGGCTCTCTCTCCCGGGCG 643  
Qy  
463 CGCTCGGGCTGACATGAGGACATCAAGCTTTCTTCCCTCTCCAGCTTCCCGGACGCTCG 522  
Db  
644 CGCGCGGGTGGACTGGTTCATCAGCTCTTCTCCCTCTCCAGCTTCCCGGACGCTCG 703  
Qy  
523 TCATGGGATCCGCTGCTCGGAGGATGATGAGGCGCTGCTGCGGCGGCGGACGCTCATGG 582  
Db  
704 TCATGGGATCCGCTGCTGATGCGCAATGACG---GCCATCTCCGCTCGCTCATGG 760  
Qy  
583 TCAGGTCTGCTGCTTCCAGTGAATCATCTGATGATGATGATGATGATGATGATGATGAT 642  
Db  
761 TCAGATCTGCTGCTTCCAGTGAATCATCTGATGATGATGATGATGATGATGATGATGAT 820  
Qy  
643 ACCGCGCGCGGCGGCTGCTTCCAGGATGATGATGATGATGATGATGATGATGATGAT 702  
Db  
821 TCGCGCGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 877  
Qy  
703 TCTCTTCCGCTGCTGATCTCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 762  
Db  
878 TGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 937  
Qy  
763 CCGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 822  
Db  
938 CCGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 997  
Qy  
823 GGTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 882  
Db  
998 GGTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1057  
Qy  
883 CGCACTCCAGACCATGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 942  
Db  
1058 GTCGTCGCGGACCAACCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1117  
Qy  
943 TGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1002  
Db  
1118 TGTGCGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1177  
Qy  
1003 ACATGCTGCGCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1062  
Db  
1178 CTTCCGAGCTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1237  
Qy  
1063 CATGCGCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1122  
Db  
1238 ACAGCACTCGGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1297  
Qy  
1123 AGAGGAAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1182  
Db  
1298 ATGCCAAGGAGCTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1357  
Qy  
1183 CGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1239  
Db  
1358 CGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1417  
Qy  
1240 GAGGCCA----- 1246  
Db  
1418 AATCCACATGGTTCATCCCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1477  
Qy  
1247 -----GGCTTACAGAGTACGGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1278  
Db  
1478 AGGAGTACCGCGAGTGGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1537  
Qy  
1279 GCAGCAGGACGAGACGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1329  
Db  
1538 CGGCAAGACGCTGGACGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1597

Qy 1330 TGTGGAAGCTGGGGTCCAACTCGACGGCGAGCTGTACCCCAAG----- 1373  
Db 1598 TGACGAAGATGGGGTTCAGCTCGACGGCGAGCTGCACCCGAGAGTCTGTCGAGCTCGAGC 1657  
Qy 1374 -----GACGACGGCGAGGGAGGGCGGCGGCTGGGATGCGCGCGGCGAGCG 1422  
Db 1658 GACCGAAGCCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1717  
Qy 1423 TGATGACGGGGCTCATCTCTCATCTGATGATGATGATGATGATGATGATGATGATGATGAT 1482  
Db 1718 TGATGACGGCTCATCTCTCATCTGATGATGATGATGATGATGATGATGATGATGATGAT 1777  
Qy 1483 ACTCCAGCTCATCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1542  
Db 1778 ACTCCAGCTCATCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1837  
Qy 1543 CAGCGATCATCGCGCTGATTTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 1602  
Db 1838 CAGCATCTGTCGAGAGTCCATCTCTCTCTGAGCGAGGCTGCTGCTGCTGCTGCTGCTGCT 1897  
Qy 1603 TCAGCTTAGGCTCTTTCATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1662  
Db 1898 TTAGCTGGGATTTTCATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1957  
Qy 1663 CGGCCATCGGATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1722  
Db 1958 CCGTCTGCTCTCATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2017  
Qy 1723 TCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1782  
Db 2018 TCGCCATCGGACTCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2077  
Qy 1783 AGGGATCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1842  
Db 2078 AAGGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2137  
Qy 1843 CAGCG 1847  
Db 2138 CAGCG 2142

AK101191 2276 bp mRNA linear PLN 24-JUL-2003  
Oryza sativa (japonica cultivar-group) cDNA clone:J033030903, full  
insert sequence.

AK101191  
AK101191.1 GI:32986400  
FLU\_CDNA; CAP trapper.  
Oryza sativa (japonica cultivar-group)  
Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Eriophytaceae; Oryza.

RESULT 12  
AK101191  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS

The Rice Full-length cDNA Consortium, National Institute of  
Agrobiological Sciences Rice Full-length cDNA Project Team,  
Kikuchi, S., Sato, K., Nagata, T., Kawagashira, N., Doi, K.,  
Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,  
Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,  
Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,  
Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN,  
Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,  
Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,  
Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y.,  
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
Yoshino, M., and Hayashizaki, Y.  
Collection, mapping, and annotation of over 28,000 cDNA clones from

TITLE

japonica rice  
 Science 301 (5631), 376-379 (2003)  
 22752273  
 12869764  
 2 (bases 1 to 2276)  
 Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayashiaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Hotta,I., Iida,Y., Iishi,Y., Ikeda,R., Imamura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kanegawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M., Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A., Mizuno,K., Murakami,K., Murata,W., Nagata,T., Nakamura,M., Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H., Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K., Shingawara,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,Y., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and Yoshimura,A.  
 Direct Submission  
 Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-839-7007).  
 This clone is one of the 28K full-length cDNA clones from japonica rice.  
 URL : <http://cdna01.dna.affrc.go.jp/cDNA/>  
 NIAS Rice Full-length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and Yamamoto,M.  
 FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y., Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M., Kodama,T., Kurotsaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J., Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K. and Murakami,K.  
 Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiraoka,T., Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J., Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N., Oka,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shingawara,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K., Yasunishi,A. and Hayaehizaki,Y.  
 Location/Qualifiers  
 1. ..2276  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultiivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /clone="J033030G03"

```
Db 1118 TCAACTCCAAAGAGGTACACATGTTCTGTTGAGCTCTAGCGCTCGCGGTGCTG 1177
Qy 1176 GAGCGCGCGCGCTGACAGTCTTCGCGCGCGG- 1208
Db 1178 GAGGCCAACTCGCGAAGCGCTCAACAGCGCGCTCCACGACTTCGCTCCGCGCG 1237
Qy 1209 -----GGGCTGACCATGCGGAGTCTGCGCCAAAGAGAGCCGAGCCCTACGACGAG 1259
Db 1238 CGCGCGGAGCGCTTCGCTGCGCGCGGCACTCCCAAAGGGGTGAGTGGCAGTGTCA 1297
Qy 1260 TAGCGGCGGAGCGACTACAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1319
Db 1298 CGCGCGGAGAGAAAGCGCGCGGAGTGGAGATCGAGAGCGGCTGAAGAGCCGCGG 1357
Qy 1320 GGCGCGAGCGTGTGGAAGCTGGGGTCCAACTCGACGGCGCAGCTGTACCCCAAGAGCAG 1379
Db 1358 GCGGGGCTGCGCGCGAAGTTCGCGGTGTCGGGTGCGCGTACGTCGCGCGAGGAAG 1417
Qy 1380 GCGGAGGG-----GAGGCGGCGCGGCTGCGGAGTGCAGCGCGCGCGGAGC 1421
Db 1418 GCGCGCGCGCGCGAGTGCAGGCGGCGGAGCGCGGAGCGGAGCGCGGCGGAGC 1477
Qy 1422 GTGATGACGCGGCTCATCTCTCATCATGCTGTGAGGAGAGCTGATCGGAAACCCCAACAC 1481
Db 1478 GTGATGACGCGGCTCATCTCTCATCATGCTGTGAGGAGAGCTCATCAGAAACCCCAAC 1537
Qy 1482 TACTCCAGCTCATCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1541
Db 1538 TACTCCAGCTCATCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1597
Qy 1542 CCAGCGATCATCCCGCGTGTGATTCGATTCCTGTCGAGCGCGGCTGTCGGGATGCCCATG 1601
Db 1598 CTTCAATTAAGGGCTCAATATCAATATGTCAGATGACAGGGCTAGGAATGCTATG 1657
Qy 1602 TTGAGCTAGGCTGTTCATGCGGCTGAGCGGAGAGATCATCGCTGCGGAGAGAGCTG 1661
Db 1658 TTGAGCTAGGCTGTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1717
Qy 1662 GCGGCGATCGCGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1721
Db 1718 GCGACATTTGCAATGCGAGTGGTTCGAGTGGTCCAGCTGTATTGACGCTACTTCC 1777
Qy 1722 ATCGCGCTGCTGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1781
Db 1778 ATTGCCATTTGCGGCTGAGGAGTACTTTCGATGTTGCGATGTTTTCAGGAGCAGCTTCA 1837
Qy 1782 CAGGGATCGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1841
Db 1838 CAGGCGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1897
Qy 1842 ACAGCG 1847
Db 1898 ACAGCG 1903
```

```
RESULT 13
AX046850
LOCUS 2293 bp DNA linear PAT 15-DEC-2000
DEFINITION Sequence 37 from Patent WO0068389.
ACCESSION AX046850
VERSION AX046850.1 GI:11876332
KEYWORDS
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
```

```
REFERENCE
1
Orozco, E.M., Weng, Z., Bruce, W.B., Cahoon, R.E. and Tao, Y.
Auxin transport proteins
Patent: WO 0068389-A 37 16-NOV-2000;
E.I. DU PONT DE NEMOURS AND COMPANY (US) ; Pioneer Hi-Bred
```

```
FEATURES
source International, Inc. (US)
Location/Qualifiers
1..2293
/organism="Triticum aestivum"
/mol_type="unassigned DNA"
/db_xref="taxon:4565"
```

## ORIGIN

```
Query Match 22.6%; Score 626.8; DB 6; Length 2293;
Best Local Similarity 64.5%; Pred. No. 1.8e-61;
Matches 1199; Conservative 0; Mismatches 467; Indels 192; Gaps 10;

Qy 164 CGCGCGCATGATCACCGCGCTGACGCTTACACGCTGTCTACGCGCGTGTGCGCGTGTGA 223
Db 67 CGTGCATGATCACCGGAGAGACATCTACAGCTGTCTGCGCGGTGTGCGCGTGTGA 126
Qy 224 CGTGCATGACGCTGCGCTACCGCTCCGCTCGCTGTGCGGATCTTCACGCGGAGCA 283
Db 127 CGTGCATGTTATGCGCTACCGGCTCGGTGCGGTGTGCGGATCTTCACGCGGAGCA 186
Qy 284 GTGCTCCGCGATCAACCGCTTCGTTGGGCTTCGCGGCTGCGGCTGCGGCTGCTTCCTTCACTT 343
Db 187 GTGCTCCGCGATCAACCGCTTCGTTGGGCTTCGCGGCTGCGGCTGCGGCTGCTTCCTTCACTT 246
Qy 344 CATCTCCACCAACGACCCCTTCGCCATGAACTGCGCTTCCTGCGCGCGGAGCAAGCTGCA 403
Db 247 CATCTCCACCAACGACCCCTTCGCCATGAACTGCGCTTCCTGCGCGCGGAGCTGCTGCA 306
Qy 404 GAGGTGCGCGCTCCTCGCGCTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCTTCCTCCCGGCGC 463
Db 307 GAGCTGCTCATCTTCGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCTTCCTCCGCTACCGGTG 366
Qy 464 GCTCGG-----GCTCGACTGAGGATCAACGCTTCCTTC 496
Db 367 CGCGCGCGCACGAGGCGCGGAGGCTCGTCTGCTGAGTCAACATCAACGCTTCCTTC 426
Qy 497 CTTCTCACGCTCCCAACAGCTGCTCATGAGGATCCCGCTGCTGCGGAGGATGACG 556
Db 427 CTTGCGAGCTGCGGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 486
Qy 557 CGCGCTGCTGCGCGCGGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 616
Db 487 CGACTTCTCG---GGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 543
Qy 617 CAGCTCATGCTCTTCTCTTCTGAGTACCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTG 676
Db 544 CAGCTCATGCTCTTCTCTTCTGAGTACCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTG 603
Qy 677 CCGGAGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 736
Db 604 CCGCGCGAGCTGCGCGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 663
Qy 737 CGCGAGGCGGAGCGTTCGAGTTCGAGCGCGGCGCGAGCGCGGCTGCGCGCGCGCGCGCGCT 796
Db 664 CAACGCGCGCGAGGCGCTGCAACGCGCGAGGCTGCGCGCGCGCGCGCGCGCTGCTGCTGCTG 723
Qy 797 CTCCTCCGCGCGCGGAGCGCGCGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 856
Db 724 CGTTCATCCGCGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCTGCTGCTG 783
Qy 857 CGCTGCGGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 916
Db 784 GATCTACCGTGGCGG-----TCCACGCGCATGACGCGCGCGCGCTGCTGCTGCTGCTG 828
Qy 917 CTTCTCGCGGCTGAGATCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
Db 829 CTTCACGCGGCTGAGATCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 888
Qy 977 CAGCTTCAACCAACCGCGCTTCTTCAACATGCTGCGG----- 1014
Db 889 CAGCTTCAACCAACCGCGCTTCTTCAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 948
Qy 1015 -----CGCGCGCGGAGGAGCGGAGGAGCGGCGGCGG-----GACGAGGAGAA 1057
```

```

Db 949 GGGCCAGCCCCCGTCCGCGAGGTGTGTGTCGCGCGGACAGGGCTCCAGCAGCAGGT 1008
QY 1058 GGGCGCATCGCGCGCGCGGAGAGACTCCCGCAGCGCGAGCCCTCGCCGTG-- 1115
Db 1009 GGCCAAACAAGTTCAAGGCGCGGAGGCGGTGCGCCCTTACCCCGCCCAACCCCGGAT 1068
QY 1116 -----CCGCGCCCAAGAGGAGGA 1132
Db 1069 GATGATCCCGCGCCACCGAAGAGGAGCTTGGGGTTCCAACTCAACTCGACRAGGA 1128
QY 1133 CTTGACATGTCTGTGTGAGCTCAGCGCTCCCGCGTGTCCAGCGCGCGCGTGCAC 1192
Db 1129 CTTGACATGTCTGTGTGAGCTCAGCGCTCCCGCGTGTGCGAGGCGCAACTCCGCAA 1188
QY 1193 ---CGTCTTCGCGCGCGCGCGTGCACATCGCGAGCTCTCGCCAAAGAGCCAGGC 1249
Db 1189 CGCGTCAACCAACCGCGCGTCCACGACTTCGCGCGCGCACCGCGCGGCGAGCCAGCC 1248
QY 1250 CTACGACAGTACGGGCGCGACGACTACAGCAGCAG-----GACGAAGAA 1294
Db 1249 ACGAGACGGCGCCACACCCAGAGCGGTGAGCGCGAGCGTGCAGCGCGTGTATGAAGAAGGA 1308
QY 1295 CGGAGCGCGCGCGCGGACAGAGGG-----GGCGAGCGTGC 1333
Db 1309 CGCCAGCAGCGCGCGGTGGAGGTGGAGATCGAGAGCGGCGATGAGAGAGCCCGCGCAC 1368
QY 1334 GAAGCTGGGGTCCAACTCGACGGGCGAGCTGTACCCCAA----- 1372
Db 1369 GGGCTGGGCGCGCAAGTTCCCGGTGTGGGTGCCCTACTAGTGCCCGCGGAGRAGGG 1428
QY 1373 ---GGAACAACGGGAGGAGGCGCGCGGTGGCGATGCCCGCGCGAGCGTGTATGAC 1429
Db 1429 CGCCGACGTGCTGGGCTGAGAGGCGCGCGCACCGATGCCCGCGCGAGCGTGTATGAC 1488
QY 1430 CGCGCTCATCTCTCATGTGTGAGGAGAGCTGATCGGAAACCCCAACACTACTCCAG 1489
Db 1489 CGCGCTCATCTCTCATGTGTGAGGAGAGCTGATCGGAAACCCCAACACTACTCCAG 1548
QY 1490 CTTATCGCGGTGCTGCTGCTGCTGCTCTATAGGTGGGCGATCGAGATGCCAGCAT 1549
Db 1549 CTTATCGCGGTGCTGCTGCTGCTGCTCTATAGGTGGGCGATCGAGATGCCAGCAT 1608
QY 1550 CATCGCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1609
Db 1609 AATCAAGGGGTCCATCATCTCTGCTGATGAGGGGTAGGGATGGTATGTTACGCTT 1668
QY 1610 AGGCTGTTTATGCGGCTGCGAGCGGAGGATCATCGCGTGCAGGAAACAGCTGGCGGCAT 1669
Db 1669 AGGTCTCTTATGCTGCTGCAACCAAGATCATCTCTTGGGGAAGTCTGTGCCACATT 1728
QY 1670 CGCGATGGGCGTCCGGTGTGTCGAGGCGCGCGGTGATGCCCGCGCTTCATCGCGGT 1729
Db 1729 TGCAATGGCAGTGAAGTTCCTGACTGGGCGCGGTGATGCCCGGACCTCAATCGCGGT 1788
QY 1730 CGGTCTCGCGCGGTCTCTCTCCATCGCATCGCTGCTGAGGCTGTCTGCTGCTGAGGGGAT 1789
Db 1789 CGGGTCTCGGGAGTGTCTCTACATGTTGCCATGTTCCAGGACGACTTCCACAGGAT 1848
QY 1790 CTTGCGGTGCTGTTGGCAAGAGATGACGGGCTTCATCCCGACATCTGTGACACAGCG 1847
Db 1849 TGTTCCTATTGTGTTTGGCAAGAGTACAAATGTCATCTCTCAAAATACTTAGCAGCG 1906

RESULT 14
BT008949
LOCUS
DEFINITION
Triticum aestivum clone wdk1c.pk008.g12.fis, full insert mRNA
sequence.
ACCESSION
BT008949
VERSION
BT008949.1 GI:32128500
KEYWORDS
Triticum aestivum (bread wheat)

```

```

ORGANISM
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
REFERENCE
1 (bases 1 to 2293)
Tingley,S.V., Wolters,P., Powell,W., Dolan,M., Miao,G.-H.,
Caraher,N.R., Hanafey,M.K. and Hainey,C.P.
Direct Submission
Submitted (20-JUN-2003) Crop Genetics, E. I. DuPont de Nemours and
Company, 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104,
USA
FEATURES
source
1..2293
/organism="Triticum aestivum"
/mol_type="mRNA"
/db_xref="taxon:4565"
/clone="wdk1c.pk008.g12.fis"
ORIGIN
Query'Match 22.6%; Score 626.8; DB 8; Length 2293;
Best Local Similarity 64.5%; Pred. No. 1.8e-61;
Matches 1199; Conservative 0; Mismatches 467; Indels 192; Gaps 10;
QY 164 CGCGGCATGATCACCGCGCTGAGCTCTTACCACTGTCTGACGGCGGTGGTCCGCTGTA 223
Db 67 GTTCGGGATGATACCGGAAGACATCTACGACGTGTGGCGGGGTGGTCCGCTGTA 126
QY 224 CTTGCGCATGACGCTGGCGTACGGCTCCGTCGCTGGTGGCGCATCTTCACGCGGACCA 283
Db 127 CTTGCGCATGTTTCATGGCGTACGGGTGCGTGGGTGGTGGGCGATCTTCACGCGGACCA 186
QY 284 GTGCTCCGGGATCAACCGCTTCTGGCGCTCTTCGCGGTGCGCTCTCTCTCCACTT 343
Db 187 GTGCTCGGCGATCAACCGCTTCTGGCGCTCTTCGCGGTGCGCTCTCTCTCTCTCTT 246
QY 344 CATCTCCACCAACGACCCCTTCGCCATGAACCTGCGCTTCTTGGCGCGCACACGCTGA 403
Db 247 CATCTCCACCAACGACCCCTTCGCCATGAACCTGCGCTTCTTGGCGCGCACCTCGCTGA 306
QY 404 GAAGTGGCGGTCTCTGGCGCTGCTGGCGGTGAGCTCCCGGCGCTCTCTCTCCCGCGCG 463
Db 307 GAAGTCTGTCATCTCTCGCGCTCTCGCGCTGCGCGTGTGGACAACAGTGTCTCTCCGCT 366
QY 464 GCTCGG-----GTCGACTGGAGCATCACGCTCTTCTC 496
Db 367 CCGCGCGGACCGAGGCGCGGAGGCTGCTGCTGGACTGGACCATCACGCTCTTCTC 426
QY 497 CTTCTCAAGCTCCCAACACAGCTCTGTATGGGCTATCCGCTGCTGCGAGGCACTGACCG 556
Db 427 CTTGGCGAGCTGCCCAACACAGCTGTGTATGGGCTATCCGCTGCTGCGCGCATGTACCG 486
QY 557 CGCGTCTCGCGCGGACGCTCATGTGTCAGTCTGCTCTCCAGTGCATCATCTGGTA 616
Db 487 CGACTTCTG-----GGTCTCTATGTGCGATCTGTGTGTCGAGAGCTCATCTGGTA 543
QY 617 CAGCTCATGCTTCTCTCTTTCAGATPACCGCGCGCGCGCTGCTCTCTGACAGATT 676
Db 544 CAGCTCATGCTTCTCTCTTTCAGATPACCGCGCGCGCGCTCATCTCCAGAGATT 603
QY 677 CCGCGAGCGGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 736
Db 604 CCGCGCGAGCTGCGGCGCGAGCATCTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 663
QY 737 CGCGAGGCGGAGCTGAGCTGAGGCGCGAGCGCGCTGCGCGCGCGCGCGCGCGCGCT 796
Db 664 CAACGGCGCGAGGCGCTGACGCGCGAGGTCGCGCGCGAGCGCGCGCGCTCCACGT 723
QY 797 CTCTCTCCGCGCGGAGCGCGCGGCGGTGCGCTCACGCTGCGCAAGTCCACAGCTC 856
Db 724 CGTCATCGCGCGGTCCGCTCGCGGTCCACCGCGCGCGCGCGCGCGCGCGCTCCCG 783
QY 857 CGCTCTCGAGGCGCGCTGCTCGCACTCGCACTCCACAGACCATGACCCCGCTGTGTCAA 916

```

Db 784 GATCTACCGTGGCGC-----TCCACGCGCATGACGCGCGCGCTCCAA 828  
Qy 917 CTTCTCCGCGTGGAGATCTACTCGTGCAGTCTGTCGCGCAACCCACCCGCGCGGTC 976  
Db 829 CTTACGCGCGTGGAGATCTACTCGTGCAGAGCTGCGCGGAGCCACCGCGAGCAGTC 888  
Qy 977 CAGCTTCAACCGCGAGCTTTCTTCAACATCTGTCGCG-----1014  
Db 889 CAGCTTCAACCGAGTCCGATCTTCTTCACTTCAACGCGGAGCAAGCTGGCTAGTCCCAA 948  
Qy 1015 -----CGCGCGCAAGCGGAGCGAGCGCGGG-----GACGAGGAGAA 1057  
Db 949 GGGCGAGCCCCCGTCCCGAGGTGTGTGTGCGCGCGAGGGGTCTGACGAGCAGGT 1008  
Qy 1058 GGGCGATCGCGCGCGCGGAGGACACTCGCGCGAGCGCGAGCGCTGCGCGTG-- 1115  
Db 1009 GGGCAACAAGTTCAAGGCGGCGAGCGCTGCGCCTACCCCGCGCCCAACCCCGGAT 1068  
Qy 1116 -----CGGGCAAGAGAGGA 1132  
Db 1069 GATGATGCGCGGCGCCACGAGAGAGGTGTGGGGTTCCAACTCAAACTCGAACAAGGA 1128  
Qy 1133 CTTGCATGCTCTGTTGAGTCTCCAGCGCTCGCCCGTGTCCGAGCGCGCGCGTGA 1192  
Db 1129 GCTGCACATGTTGTTGTGAGTCTCCAGCGGTGCGCCGTTGCGAGGGCAACTCCGCA 1188  
Qy 1193 ---CGTCTTCCGCGCGCGCGTGCATGCGGAGTCTTCCGCAAGAGCGCGC 1249  
Db 1189 CGCGTCAACACGCGCGCTCCAGCGCTTCCGCGCGCGCGCGCGCGCGCGCGCGCG 1248  
Qy 1250 CTACGAGGAGTACGCGCGGCGAGCTACAGCAGCAG-----GACGAAGAA 1294  
Db 1249 ACAGAGCGCGCCACACCCAGAGCGTGTAGCGCGAGCGTGTGATGAGAGGA 1308  
Qy 1295 CGGGAGCGCGCGCGGCAAGGGCG-----GGCCGAGCGTGTG 1333  
Db 1309 CGCCAGCAGCGCGCGTGGAGTGGAGATCGAGGAGCGCGATGATGAGAGCGCGCGAC 1368  
Qy 1334 GAAGCTGGGTCCAACTCGAGCGCGAGTGTACCCCA-----1372  
Db 1369 GGGCTGGGCGCAAGTTCCGTTGCGGGTCCCGCTACGTGGCGCGCGGAGAGGG 1428  
Qy 1373 ---GGAAGAGCGGAGGCGCGCGTGGCGATGCCCGCGCGAGCGTGTATGAC 1429  
Db 1429 CGCGGAGCGTGGCTGGAGGAGCGCGCACCCGATGCCCGCGCGCGAGCGTGTATGAC 1488  
Qy 1430 GCGGCTCATCTCATGTTGGAGGAGTGTATCGGAGACCCCAACACTTCTCCAG 1489  
Db 1489 CCGGCTCATCTCATGTTGGCGAGCTCATCCGCAACCCCAACACTTCTCCAG 1548  
Qy 1490 CCTCATCGCGCTGCTGCTGCTCTCTTACAGGTGGGCGATCGAGATGCCAGCGAT 1549  
Db 1549 CCTCATCGCGCTGCTGCTGCTCACTGCTCTCTTACAGGTGGGCAATTCAGTGCCTACAAT 1608  
Qy 1550 CATCGCGCGTGGATTCGATCTGTCGAGCGGGTCTCGGATGGCGCATGTTGAGCT 1609  
Db 1609 AATCAAGGGTTCATATCCATCTGCTGATGAGGGGTAGGATGGCTATGTTGAGCTT 1668  
Qy 1610 AGGCTCTTTCATGGCGCTGCGAGCGGAGGATCATCGCGTGGGGAACAAGCTGGCGGCCAT 1669  
Db 1669 AGTCTCTTTCATGGCTGCAACCAAGATCATCTTGGGGAGTCTGTGCGCACATT 1728  
Qy 1670 CGGATGGGCGTCCGTTGTGTCAGCGCGCGCGGTCATGCGCGCGCTTCCATCGCGT 1729  
Db 1729 TGCAATGGCAGTGGGTTCTTGACTGGCGCGCGGTGATCGCGCGACCTCAATCGCGGT 1788  
Qy 1730 CGGTCTGCGCGGTCTCTCTCCATCGCATGCTCCAGGTGCTCTGCTCAGGGAT 1789  
Db 1789 CGGGCTCGGGGAGTGTCTCATATGTTGCCATGTTCCAGGAGAGACTTCCACAGGAAT 1848  
Qy 1790 CGTGGCGTTCGCTTCCGCAAGAGTACGGCGTTTCATCCGACATCTTGAAGCAGCG 1847  
Db 1849 TGTTCATTTGTGTTCGCAAGGAGTAAATTGCCATCTCAAACTTATAGCAGCG 1906

RESULT 15  
AX654121  
LOCUS AX654121 1827 bp DNA linear PAT 22-MAR-2003  
DEFINITION Sequence 3991 from Patent WO03000898.  
ACCESSION AX654121  
VERSION AX654121.1 GI:29156935  
KEYWORDS  
SOURCE Oryza sativa  
ORGANISM Oryza sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.  
1  
REFERENCE Chang H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,  
AUTHORS Katagiri, F., Qian, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.  
TITLE ' Plant genes involved in defense against pathogens  
JOURNAL Patent: WO 03000898-A 3991 03-JAN-2003;  
SYNOPSIS Syngenta Participations AG (CH)  
FEATURES Location/Qualifiers  
1 .1827  
/organism="Oryza sativa"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:4530"  
ORIGIN  
Query Match 22.3%; Score 618; DB 6; Length 1827;  
Best Local Similarity 63.2%; Pred. No. 1.9e-60;  
Matches 1116; Conservative 2; Mismatches 522; Indels 126; Gaps 6;  
Qy 202 TGACGGCGGTGGTGCCTGTAGTGGCCATGACGCTGGCGTACGGCTCCGTCGGTGGT 261  
Db 2 TGGCGGCGGTGGTGCCTGTAGTGGCGATGTTCTTGGGTACGGGTGGTGGTGGT 61  
Qy 262 GGGCGATCTTACGCGCGGACAGTGTTCGGGATCAACCGCTTCGTGGCGCTCTTGGCG 321  
Db 62 GGGCGATCTTACGCGCGGACAGTGTTCGGGATCAACCGCTTCGTGGCGCTCTTGGCG 121  
Qy 322 TGGCGCT 381  
Db 122 TGGCGCT 181  
Qy 382 TCTTGGCGCGGACACGCTGCGAGAGGTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 441  
Db 182 TCTTGGCGCGGACACGCTGCGAGAGGTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 241  
Qy 442 GCGGCT 501  
Db 242 GCTTCT 301  
Qy 502 CCAGCTTCCCAACACGCTGCTCATGGGCATCCGCTGCTGCGAGGATGTACGGCGCT 561  
Db 302 CCAGCTTCCCAACACGCTGCTCATGGGCATCCGCTGCTGCGAGGATGTACGGCGCT 358  
Qy 562 GGTGGCGGACAGCTCATGCTCCAGTCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 621  
Db 359 CATACTCCGCT 418  
Qy 622 TCATGCTCTTCT 681  
Db 419 TGATGCTCTTCT 478  
Qy 682 ACGGCGCGCGCGGTCTCATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 741  
Db 479 A---CAGCGCGCGGTCTCATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 535  
Qy 742 GGGGGGACGTCGAGCTCGAGCGCGAGCGCGCGCTGCGCGCGCGCGCGCGCGCGCTCTCT 801  
Db 536 GCGGCGAGCGGAGCGGAGCGCGGAGGTGGCGCGGAGCGCGCGCGCTGCGCTCACCGTGC 595  
Qy 802 CCCGCGCGCGGAGCGCGCGCGGTGCGCGCTCACCGTGGCGAAGTCCACGAGCTCGGCT 861



596 GCCGTCCTCGTGTGCGCGGGTCGCTGCTCAAGCGCGCGCGCTCGAAGCTGACGG 655  
 862 CCGAGCGCGTGTCTCGCACTCGCACTCCAGACCATGACGCCCGTGTGTCTCAACCTCT 921  
 656 GAGCGGAGATCTACTCGCTTAGCTGTGTGCGGAAACCAACCCGCGGGGTCTCAACTTCA 715  
 922 CCGCGGTGAGATCTACTCGCTGAGTGTGCGCAACCCACCGCGCGGGTCTCAGCT 981  
 716 ACCACCGCGACTTCTTCGGCAATGTTGCGCGCGGGCCACCGCCCCCGACGCCCTCGG 775  
 982 TCAACCAACCGCACTTCTTCAACATCTGTGCGCGCGCGCCCAAGGGAGCGGAGAGCGG 1041  
 776 TCGCGGCTCGAGCTTCGGGGCTCGAGCTCTACTCGCTCAATCGTCTCGGGGCCCAA 835  
 1042 CCGGGGACGAGGAGAGGGCGCATGCGCGCGCGCGCGCGGAGAGACACTCGCGCGAGCCG 1101  
 836 CCGGAGGAGTCTCAACTTTCGAGAGCACTCGGCACCGCCCGAAACCAACCGGCAACGA 895  
 1102 AGCCGCTCGCGGTGCGCGCAAGAGAGAGACTTCACATCTCGTTGGAGTTCACAGCG 1161  
 896 CCACGGGGGCACTCAACACAGATGCCNAGGAGTTCACATGTTGTTGTGAGCTCGAGCG 955  
 1162 CTTGCGCGGTGTCGAGCGCGCGCGGTGACGCTTC---GGCGCGCGCGCGCTGACC 1218  
 956 CGTCTCCGCTCTCAGAAATCAGCGGCTCGCTGTGTTCACTGCGCGCGCGCGCGCGCG 1015  
 1219 ATCGCGAGCTCTCGCCCAAGAGGCCA----- 1246  
 1016 CTCTGACGTGCGCGCAAGGAATCCATGCTCATCCCGCGCACTCGCGCGAGAACA 1075  
 1247 -----GGCCTACGACG 1257  
 1076 ACGGCTCAGCAAGAGCACGAGGAGTACGGCGCAGTGGCATTTGGTGGCGGGCGCGG 1135  
 1258 AGTACGGGCGGACGACTACAGCAGCAGCAGCAAGAACGGGAGCGGGCGCGGACAAGG 1317  
 1136 GAGAGAACTTCACTTCGAGGCGGCAAGACGTTGACGCGCGCGGAGGAGTACGAGG 1195  
 1318 GCGGGCC-----GACGCTGTGGAAGCTGGGTCCAACTCGACGCGCGCAGCTGTACC 1368  
 1196 AGCGGGCTTGCTGACGGGCTGACGAGATGGGTTCGAGCTCGACGCGGAGCTGCACC 1255  
 1369 CCAAG-----GACGACGGCGAGGGAGGGCGCGCGG 1401  
 1256 CGAAGTCTGTCGACGTCGACGACGAAACGCGCGCGCGCGCGCGCGCGCGAGT 1315  
 1402 TGCGGATGCGCGCGGAGCGTGATGACGCGCTCATCTCATGTTGTGGAGGAGC 1461  
 1316 ACCAATGCGCGCGGACGCTGATGACACGCTCATCTCATTAATGTTGTGGCGCAAGC 1375  
 1462 TGATCCGGAACCCAAACCTACTCCAGCTCATCGGCTGCTGTGTTCCCTGCTCCT 1521  
 1376 TCATCGCAACCCCAACTTACTCCAGCTCTCGGCTCGCTGCTGCTGCTGCTGCT 1435  
 1522 ACAGGTGGGACTCGAGATGCCAGCGATCATGCCCGCTCGATTTGATCTCTCGGACG 1581  
 1436 TCGGTGGGACGCTCCATGCCAGCAATCGTCGAGAGTCCATCTCCATTTCTCTGGAGC 1495  
 1582 CGGCTCTCGGATGGCCATGTTTCAGCTAGGCTGTTTCATGGCGCTGCGAGCCGAGATCA 1641  
 1496 CAGGCTGGGATGGCCATGTTTAGCTGGGATGTTTCAATGGCGTGCAGCCCAAGATCA 1555  
 1642 TCGGTGCGGGAACAAGCTGGCGGCGCATGCGATGGGCTCGGTTCGTTGTCAGGCCCG 1701  
 1556 TCGCGTGTGGCAATCAGCGCGCGCTGCTCCATGCGCGCTCGCTTCTCGCGGGCCGTG 1615  
 1702 CGGTTCATGGCGCGGCTCATCGCGCTCGGTCTCGCGCGGCTCTCTCCACATCGCCA 1761  
 1616 CCGTTCATGGCGCGGCTCAATCGCCATCGGACTCCGCGGAGCGCTCTGCACTGCCA 1675  
 1762 TCGTTCAGGCTGCTCTGCTCTCAGGGAGTCGTGCGCTTCTGTTGCGCAAGGAGTACGGCG 1821  
 1676 TTGTTGAGGCGGCTTACCAAGGAGTGTGCTTTGTTTGTGAAAAGAAATACAAATG 1735



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 3, 2004, 00:50:45 ; Search time 1036 Seconds  
(without alignments)  
11354.499 Million cell updates/sec

Title: US-10-030-884-13  
Perfect score: 2769  
Sequence: 1 ccacgcctccgcctgcgcct.....aaaaaaaaaaaaaaaaaa 2769

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

```
Database :
N_Geneseq 25Jan04:
1: Geneseqn1980s:
2: Geneseqn1990s:
3: Geneseqn2000s:
4: Geneseqn2001as:
5: Geneseqn2001bs:
6: Geneseqn2001cs:
7: Geneseqn2003as:
8: Geneseqn2003bs:
9: Geneseqn2003cs:
10: Geneseqn2004s:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query			ID	Description
	Score	Match	Length		
1	2769	100.0	2769	3	AAa94721
2	838.6	30.3	1926	7	ADa70391
3	760	27.4	1686	7	ADa69470
4	683	24.7	1984	7	ADa70233
5	671.2	24.2	2374	3	AAc22980
6	626.8	22.6	2293	3	AAc94733
7	618	22.3	1827	3	ADa70668
8	563	20.3	1845	7	ADa69396
9	536.2	19.4	2162	3	AAa94724
10	452.6	16.3	504	3	AAa94720
11	425.6	15.4	1426	3	AAa94719
12	347.2	12.5	1851	3	AAc43229
13	329.8	11.9	2292	2	AAV16361
14	323.8	11.7	1088	3	AAa94716
15	319	11.5	1860	6	AB213043
16	319	11.5	1860	7	ADa68467
17	319	11.5	2027	3	AAc44892
18	316.6	11.4	2549	3	AAa94729
19	305.6	11.0	1713	3	AAc48532
20	301.8	10.9	2101	3	AAa94727
21	290.8	10.5	1923	3	AAc44722
22	290.8	10.5	1923	6	AB213042
23	267.6	9.7	1725	3	AAc46518

## ALIGNMENTS

## RESULT 1

AAA94721

—

100

[illegible]

XX  
ECC

KW plant breeding: herbicide: ss:  
KW plant breeding: herbicide: ss:  
KW plant breeding: herbicide: ss:

Zea mays.

AA PN WO200068389-A2.

16-NOV-2000

03-MAY-2000: 2000WO-IIS012061-XX PF

XX  
PR 07-MAY-1999. 99IIS-0133040P.XX  
BA  
(DUBO) DU PONT DE NEMOURS & CO ET TPA (PION-) PIONEER HI-BRED INT. INC.  
XX

PL Orozco EM, Meng Z, Bruce WB, Calhoun KE, Tao L, YY

DR WPI; 2000-687647/67.  
DR B-DCDB: AAB26934

XX BT New Mexico and

PT for modulating root growth of plants and to screen for herbicides.  
yy

PS Claim 2; Page 60-61; 94pp; English.

Auxins are plant hormones that initiate

responses and organ (e.g. flower/leaf) development. The present invention relates to corn auxin transport protein coding sequences and proteins. The present sequence is one such coding sequence. This sequence may be used to modulate root development, e.g. to produce a more robust root system, alter root angle or redirect root growth. Also, the present sequence may be useful for gene mapping (e.g. for plant breeding) and to identify loss of function mutants. The protein encoded by the present sequence may be useful for raising specific antibodies, for the detection

CC	of auxin transport proteins and to design and/or identify specific
XX	inhibitors of auxin transport proteins, potentially useful as herbicides
SQ	Sequence 2769 BP; 527 A; 917 C; 757 G; 568 T; 0 U; 0 Other;
	Query Match 100.0%; Score 2769; DB 3; Length 2769;
	Best Local Similarity 100.0%; Pred. No. 0;
	Matches 2769; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 CCACGGCTCCGTGAGCCCTCAACACCTCTCTTCCTTCATGTCACACTACCATCTCA 60
Db	1 CCACGGCTCCGTGAGCCCTCAACACCTCTCTTCCTTCATGTCACACTACCATCTCA 60
Qy	61 TCTCTCCGCAATTTTACACCACCTCCCTCTCTGTTGCAACCAACAATTTGGCACTGCTCG 120
Db	61 TCTCTCCGCAATTTTACACCACCTCCCTCTCTGTTGCAACCAACAATTTGGCACTGCTCG 120
Qy	121 TGC CGACCCCTCTCTCCCTTCGCCGCTCCCCCGCAAGAAGCATCGCGGCATGATCACCG 180
Db	121 TGC CGACCCCTCTCTCCCTTCGCCGCTCCCCCGCAAGAAGCATCGCGGCATGATCACCG 180
Qy	181 CGGTGACCTCTACCAAGCTGCTGACGGCGGTGGCGCTGTACGTGGCCATGACGCTGG 240
Db	181 CGGTGACCTCTACCAAGCTGCTGACGGCGGTGGCGCTGTGACGTGGCCATGACGCTGG 240
Qy	241 CGTACGGCTCCGTCGCGCTGGTGGCGCATCTTTACGCCCGACCAAGTGTCCGGGATCAAC 300
Db	241 CGTACGGCTCCGTCGCGCTGGTGGCGCATCTTTACGCCCGACCAAGTGTCCGGGATCAAC 300
Qy	301 GCTTCTGGCGCTCTTCGCGGTGGCGCTCTCTCTTCCTTCCTTCATCTCCACCAAGCAC 360
Db	301 GCTTCTGGCGCTCTTCGCGGTGGCGCTCTCTCTCTTCCTTCCTTCATCTCCACCAAGCAC 360
Qy	361 CTTTCGCCATGAACCTGGCGCTTCTTGGCGCGCGCAACGCTGCAAGTGGCGCTCTCGT 420
Db	361 CTTTCGCCATGAACCTGGCGCTTCTTGGCGCGCGCAACGCTGCAAGTGGCGCTCTCGT 420
Qy	421 CGCTGTGGCGCTGGCGCTCCGGCGCTCTCTCTCCCGCGCGCTCGGCTCGACTGCA 480
Db	421 CGCTGTGGCGCTGGCGCTCCGGCGCTCTCTCTCCCGCGCGCTCGGCTCGACTGCA 480
Qy	481 GCATCAGCTCTTTCTCCCTCTCCAGCTCCCAACACGCTCGTCAATGGGCATCCCGCTCG 540
Db	481 GCATCAGCTCTTTCTCCCTCTCCAGCTCCCAACACGCTCGTCAATGGGCATCCCGCTCG 540
Qy	541 TGGAGGCAATGACGGCGGCTCGTTCGGCGGCAAGCTCATGGTCCAGTGGTGGTCTCTCC 600
Db	541 TGGAGGCAATGACGGCGGCTCGTTCGGCGGCAAGCTCATGGTCCAGTGGTGGTCTCTCC 600
Qy	601 AGTGCATCATCTGGTACAGCTCATGCTCTTCTCTTCAGATACCGCGCGCGCGCGCGCG 660
Db	601 AGTGCATCATCTGGTACAGCTCATGCTCTTCTCTTCAGATACCGCGCGCGCGCGCGCG 660
Qy	661 TCGTCTCGACAGTTCGCCGAGCGCGCGTGCATGCTCTCTTCGGGTGCTGACT 720
Db	661 TCGTCTCGACAGTTCGCCGAGCGCGCGTGCATGCTCTCTTCGGGTGCTGACT 720
Qy	721 CCGAGCTCGTCTCGCTCGCCAGGGGAGCTCGAGCTCGAGCGCGAGCCGAGCGGCGTGC 780
Db	721 CCGAGCTCGTCTCGCTCGCCAGGGGAGCTCGAGCTCGAGCGCGAGCCGAGCGGCGTGC 780
Qy	781 CCGCGCGCGCGCGCTCTCTCCCGCGCGGGAGCGCGCGCGGTGCGGTCAACGTGC 840
Db	781 CCGCGCGCGCGCGCTCTCTCCCGCGCGGGAGCGCGCGCGGTGCGGTCAACGTGC 840
Qy	841 GCAAGTCCACAGCTCGCGCTCCGAGGCGCGTGTCTCGCACTCCAGACCATGC 900
Db	841 GCAAGTCCACAGCTCGCGCTCCGAGGCGCGTGTCTCGCACTCCAGACCATGC 900
Qy	901 AGCCCCGTGTCCAACTCTCCGGCGTGGAGATCTACTCGCTGAGTGTGTCGGCGAAC 960
Db	901 AGCCCCGTGTCCAACTCTCCGGCGTGGAGATCTACTCGCTGAGTGTGTCGGCGAAC 960

Qy	961	CAACCCCGCGCGGGTCCAGCTTCAACCAAGCGGACCTTCTTCAACATCTCGGCGCGCGCG	1021
Db	961	CCACCCCGCGCGGGTCCAGCTTCAACCAAGCGGACCTTCTTCAACATCTCGGCGCGCGCG	1020
Qy	1021	CCAAGGAGCGGAGGAGCGCGGGGGACGAGAGAGAGGCGCGATCGCGCGCGCGCGCGCG	1080
Db	1021	CCAAGGAGCGGAGGAGCGCGGGGGACGAGAGAGAGGCGCGATCGCGCGCGCGCGCGCG	1080
Qy	1081	GAGGACACTCGCGCAGCGCAGCGCGCGTCCGCGTCCGCGCAAGAGAGAGAGACCTGCACA	1140
Db	1081	GAGGACACTCGCGCAGCGCAGCGCGCGTCCGCGTCCGCGCAAGAGAGAGAGACCTGCACA	1140
Qy	1141	TGTCGTCGTGAGACTCCAGCGCCTCGCCCGTGTCCGAGCGCGCGCGCTGTGACGCTCTTCG	1200
Db	1141	TGTCGTCGTGAGACTCCAGCGCCTCGCCCGTGTCCGAGCGCGCGCGCTGTGACGCTCTTCG	1200
Qy	1201	GCSCCGCGCGCTGACCATCGCAGCGTCTCGCCAAAGAGAGCGCCAGGCGCTACGACGAGT	1260
Db	1201	GCSCCGCGCGCTGACCATCGCAGCGTCTCGCCAAAGAGAGCGCCAGGCGCTACGACGAGT	1260
Qy	1261	ACGGGCGCGACGACTACGAGCAGACGAGAAACGGGAGCGCGCGCGCGGACAAAGGCGG	1320
Db	1261	ACGGGCGCGACGACTACGAGCAGACGAGAAACGGGAGCGCGCGCGCGGACAAAGGCGG	1320
Qy	1321	GGCCGACGCTGTCGAAAGCTGGGTCGAACTCGACGCGCAGCTGTATCCCAAGAGACGACG	1380
Db	1321	GGCCGACGCTGTCGAAAGCTGGGTCGAACTCGACGCGCAGCTGTATCCCAAGAGACGACG	1380
Qy	1381	GCAGGGAGGGCGCGCGGGTGGCGATGCCCGCGCGAGCGGTGATGACCGGCTCATTC	1440
Db	1381	GCAGGGAGGGCGCGCGGGTGGCGATGCCCGCGCGAGCGGTGATGACCGGCTCATTC	1440
Qy	1441	TCATCATGTTGTGAGGAGCTCATCCGGAACCCCAACACTACTCCAGCGCTCATCGCGG	1500
Db	1441	TCATCATGTTGTGAGGAGCTCATCCGGAACCCCAACACTACTCCAGCGCTCATCGCGG	1500
Qy	1501	TCGTCGTGTCCTGGTCTCCTACAGTGGGGCATTCGAGATGCCAGCGATCATTCGCCCGGT	1560
Db	1501	TCGTCGTGTCCTGGTCTCCTACAGTGGGGCATTCGAGATGCCAGCGATCATTCGCCCGGT	1560
Qy	1561	CGATTTCCATCTCTCGAGCGCGGCTCTCGGATGGCCATGTTTCAGCCTTAGGCGCTGTCA	1620
Db	1561	CGATTTCCATCTCTCGAGCGCGGCTCTCGGATGGCCATGTTTCAGCCTTAGGCGCTGTCA	1620
Qy	1621	TGGCGCTGCAGCGAGGATCATCGCGTGGGGAACAGCTGGCGGCCCATCGCGATGGGCG	1680
Db	1621	TGGCGCTGCAGCGAGGATCATCGCGTGGGGAACAGCTGGCGGCCCATCGCGATGGGCG	1680
Qy	1681	TCGCGTTCGTCGAGGCGCGCGTCAATGCGCGCGCGCTCCATCGCGCGTCTCGTTCGCGG	1740
Db	1681	TCGCGTTCGTCGAGGCGCGCGTCAATGCGCGCGCGCTCCATCGCGCGTCTCGTTCGCGG	1740
Qy	1741	GCCTCTCTCCACATCGCGCATGCTCAGGCTGCTGCGCTCAGGGGATCGTGCCTTCG	1800
Db	1741	GCCTCTCTCCACATCGCGCATGCTCAGGCTGCTGCGCTCAGGGGATCGTGCCTTCG	1800
Qy	1801	TGTTCCGCAAGGAGTACGGCGTTTCATCCCGACATCTCTGAGCACAGCGGTATGGTCCAATAA	1860
Db	1801	TGTTCCGCAAGGAGTACGGCGTTTCATCCCGACATCTCTGAGCACAGCGGTATGGTCCAATAA	1860
Qy	1861	CATCGCATGTTTTCATCAGCTTGCATAGTTTAAACGGGAAAAAAGACAGAGCAATCCGATG	1920
Db	1861	CATCGCATGTTTTCATCAGCTTGCATAGTTTAAACGGGAAAAAAGACAGAGCAATCCGATG	1920
Qy	1921	ACGACGCACTGAATTTCACTATGATTTCACTAATGATGGTGTGTTTCATGCAAGTGCAGTC	1980
Db	1921	ACGACGCACTGAATTTCACTATGATTTCACTAATGATGGTGTGTTTCATGCAAGTGCAGTC	1980
Qy	1981	AAAGAACACTAATTAAGCACTGATCTAGGACGATCAGCATGATTCATTTGCTTTGTTTC	2040
Db	1981	AAAGAACACTAATTAAGCACTGATCTAGGACGATCAGCATGATTCATTTGCTTTGTTTC	2040
Qy	2041	TCCTGCACAATCTGCAATTTCTTACTACACAGTGTGCCCTTCACTCATCCATCAGATGATCA	2100







Db 2 TGGTGCCGCTGTACGTGGCGATGATCTTGGCGTACGGGTGGTGAAGTGGTGGCGCATCT 61  
Qy 271 TCACGCGGACAGTCTCCGGGATCAACCGCTTCGTGGCGCTCTTCGCGTCC 330  
Db 62 TCACGCGGACAGTCTCCGGGATCAACCGCTTCGTGGCGCTCTTCGCGTCC 121  
Qy 331 TCTCTTCCATCTCATCTCAACAGACCCCTTCGCCATGAACCTTCGCTTCCTGGCG 390  
Db 122 TCTCTTCCATCTCATCTCAACAGACCCCTTCGCCATGAACCTTCGCTTCCTTCGCG 181  
Qy 391 CGACACGCTGCAGAGGTGGCGTCTCGCGCTGTGGCGTGGCGTCCGCGGCTCT 450  
Db 182 CGACACCTCTGAGAGCTCATCTGCTCTCGCGTCTCTCTCTCTCTCTCTCTCTCTCT 450  
Qy 451 CTTCCCGCGCGGCTCGGCTCGAGTGGAGCATCAAGCTCTCTCTCTCTCTCTCTCTCTCTCT 510  
Db 233 ACTCTTCGCGCGGCT 292  
Qy 511 CCAACACGCTCTCATGGGATCCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 570  
Db 293 CCAACACGCTCTCATGGGATCCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 349  
Qy 571 GCACGCTCATGGTCCAGGTCT 630  
Db 350 GTAGCCTCATGGTCCAGGTCT 409  
Qy 631 TCCCTCTTCAGTACCGCGCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 690  
Db 410 TCATGTTCAGTACCGCGCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 469  
Qy 691 CCGCGTCCATCT 750  
Db 470 GCG---CCATCGGTCT 526  
Qy 751 TCAGTCTGAGGCGCGAGCGCGGTCT 810  
Db 527 ACATGATCAGACGGAGCGAGGTGAA----- 554  
Qy 811 GCGACCGCGCGGTCT 870  
Db 555 -GGAGACGGCAGATACAGTCACTGCGTGGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 613  
Qy 871 CGTGCTCGCACTC---GCACTCCAGACCATGAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 927  
Db 614 ACTCGCGGCTCTCATGGGTCTCTCCAGCACACCGCGCGCGCGAGCACTCTCAACAG 673  
Qy 928 TGAGATCTACTCTGCTGAGTCTGCTGCGCAACCCACCGCGCGCGGTCTCAAGTCAACC 987  
Db 674 CCGAGATCTACTCTGCTGAGTCTGCTGCGCAACCCACCGCGCGCGGTCTCAAGTCAACC 733  
Qy 988 ACGCCGACTTCTTCAACATCTCTCGCGCGCGCGCGCGCGAGGAGCGCGCGCGG 1047  
Db 734 ACACTGACTTCTACT 793  
Qy 1048 ACAGAGAGAGGCGC----- 1063  
Db 794 GGTGCGCAACCGCGCGCAAGCGCGCGCTCTCAACTACAGAGAGACGCGCGCGCGCA 853  
Qy 1064 ----ATGCGCGCGCGCGCGAGGACACTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1119  
Db 854 ACAAGCGCGCGCGCGCGAGTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 913  
Qy 1120 CCAAG-----AGGAAGGACCTTGCA 1140  
Db 914 CCAAGCG 973  
Qy 1141 TGGTCTCTGAGCTCCAGCGCT 1200  
Db 974 TGTCTCTGAGCTCCAGCGCT 1033  
Qy 1201 GCGCGCGCGCGCTGACCATGCGAGCT 1260  
Db 1034 ACAACGACCGCGCGCGCTCAAGAGGTCTCGCATGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1093

Qy 1261 ACGGGCGCGACGACTACAGCAGCAGCAAGAACCGGAGCGCGCGCGCGCGCAAGGGCG 1320  
Db 1094 ACGGCTGGAGAGGAGCGACTTCAGCTTCGGGAACAGAGCGCTCGCCGAGAGGAGCGCG 1153  
Qy 1321 GCGCGAGCTGTGGAAGCTGGGTCAACTCAGCGCGCGAGCTGTACCCCAAGAGGAGCG 1380  
Db 1154 AGGCGCGCGAGAGAGCGTGGCGCGCGGTGTGGGTGAGCATGGCAAGCCCTGGGT 1213  
Qy 1381 GCGAGGCGGCGCGCGCGGTGGCGATCGCGCGCGCGAGCGGTGATGACCGGCTCATCC 1440  
Db 1214 T-----GACCGCGCGCGCGCGCGATCGCGCGCGAGCGGTGATGACCGGCTCATCC 1267  
Qy 1441 TCATCATGGTGGAGAGCTGATCCGAAACCCCAACCTACTCTCAGCTCAGCTCATCGCG 1500  
Db 1268 TGATCATGGTGTGGCGAAGCTCATCCGAAACCCCAACCTACTCTCAGCTCAGCTCATCGCG 1327  
Qy 1501 TCGTCTGCTCCCTGCTCTCTACAGTGGGCGATCGAGATGCCAGCGATCATCGCGCGGT 1560  
Db 1328 TCATCTGCTGCT 1387  
Qy 1561 CGATTTCATCTCTGTCGAGCGCGGTCTCGGATGCGGATGCGGATGCGGATGCGGATGCGG 1620  
Db 1388 CCATCTCATCTCTCTCGAGCGCGCTCGGATGCGGATGCGGATGCGGATGCGGATGCGG 1447  
Qy 1621 TGGCGTGCAGCGAGGATCATCGTGGGGAACAAAGCTGGGCGCATCGCGATGCGG 1680  
Db 1448 TGGCGTGCAGCGCGGATCATCGTGGGGAACAAAGCTGGGCGCATCGGATGCGG 1507  
Qy 1681 TCGGTTCTGTCGAGGCGCGCGGTCTCATGCGCGCGCTCCATCGCGCTGCTCTCGCG 1740  
Db 1508 TGGGTTCTCTGACCGCGCGCGGTCTCATGCGCGCGCTCCATCGCGCTGCTCTCGCG 1567  
Qy 1741 GCGTCTCTCTCCATCGCGCATGTCAGGCTCTCTGCTCAGGGGATGCTGCGCTTCTG 1800  
Db 1568 GCAGCTCTCTCCAGCTCGCATGTCGAGGAGCGTTGCMAGGAGCATGTCCTCTG 1627  
Qy 1801 TGTTCGCAAGGAGTACGGGTTCATCCGACATCTCTGAGCAGCG 1847  
Db 1628 TCTTCGCAAGGAGTACAGGTGCACCGCGACATCTCTCAGCAGCG 1674

## RESULT 5

AAZ29280  
ID AAZ29280 standard; cDNA; 2374 BP.

AC AAZ29280;

DT 28-FEB-2000 (first entry)

DE Rice EIR1 Homologue (REH1) cDNA.

Root-specific auxin transport protein; EIR1; Ethylene Insensitive Root;  
gravitropism; auxin-based herbicide; plant root; transgenic plant;  
auxin homeostasis; Rice EIR1 Homologue; REH1; ds.

OS Oryza sativa.

Key Location/Qualifiers  
158..1945  
/\*tag= a  
/product= "REH1 protein"  
FT sig\_peptide 158..337  
/\*tag= b

PN WO9963092-A1.

PD 09-DEC-1999.

PF 03-JUN-1999; 99WO-US012277.

PR 03-JUN-1998; 98US-0087789P.



PA (WHED) WHITEHEAD INST BIOMEDICAL RES.  
XX Luschnig C, Gaxiola RA, Grisafi P, Fink GR;  
XX WPI; 2000-086979/07.  
DR P-PSDB; AAY44265.  
XX  
PT DNA encoding a root-specific auxin transport protein, used to develop  
PT transgenic plants with increased resistance to herbicides.  
XX  
PS Claim 1; Fig 8; 55pp; English.  
XX  
CC The present sequence is a cDNA encoding REH1, a rice homologue of EIR1  
CC which functions as a root-specific auxin transport (efflux) carrier  
CC protein involved in gravitropism. The sequence is obtained from a rice  
CC EST derived from root-specific cDNA. The sequence is used for producing  
CC genetically engineered plants with greater resistance to auxin-based  
CC herbicides and auxin transport inhibitors in combination with a second  
CC herbicide. It can also be used to enhance transport of auxin in plant  
CC roots, produce transgenic plants which exhibit altered auxin homeostasis  
CC and mutant plants in which the roots are agravitropic and have reduced  
CC sensitivity to ethylene  
XX  
SQ Sequence 2374 BP; 471 A; 736 C; 733 G; 434 T; 0 U; 0 Other;  
Query Match 24.2%; Score 671.2; DB 3; Length 2374;  
Best Local Similarity 66.8%; Pred. NO. 3e-98;  
Matches 1196; Conservative 0; Mismatches 423; Indels 171; Gaps 10;  
QY 165 GGGCCATGATCACCGCTGTGACCTCTACACGCTGTGACGGGTGGTCCGCTGTAC 224  
DB 152 GCGAAGATGATTACCGCGCGGACCTTACACAGTATGACGGGATGGTCCGTTGTAC 211  
QY 225 GTGGCCATGACCTGGCTACGGCTCCGCTGCTGGTGGGCGATCTTCAAGCCGGAACG 284  
DB 212 GTGGCGATGATCTGGCTACGGGTGCGGTGAAGTGGTGGCGATCTTCAAGCCGGAACG 271  
QY 285 TGCTCCGGGATCAACCGCTCTGTGGCGCTCTTCCGCGTGGCGCTCTCTCCCTCCATTC 344  
DB 272 TGCTCCGGGATCAACCGCTTGTGGCGCTCTTGGCGGTGGCGCTGTCTGTCTTCACTTC 331  
QY 345 ATCTCCACCAACGACCTCTCGCCATGAACCTTGGCTCTTGGCGCGGACACGCTCGAC 404  
DB 332 ATCTCCACCAACGACCTTACAGATGAACCTTCCGGTTTCATCGCGCGGACACGCTCGAC 391  
QY 405 AAGTGGCGCTCTCGCGCTGTGGCGCTTGGCTTCCGCGGCTCTCTCTCCCGCGCGG 464  
DB 392 AAGCTGATGGTGTGGCTGCTCATCGCGCTGAGACCACTCAGC-----CGCGG 442  
QY 465 CTCGGGCTCGACTGAGCATCAAGCTCTTCTCCCTCTCCAGCTTCCCAACACGCTCGTC 524  
DB 443 GGGAGCTCGAGTGAGACCATCAAGCTCTTCTCCCTCTCCAGCTTCCCAACACGCTCGTC 502  
QY 525 ATGGGATCCGCTCTCGGAGGATGTACGGCGGTGCTCGCGGCGGACAGCTCATGCTC 584  
DB 503 ATGGGATCCCTTTGCTCAAGGGCATGTACGGGAGTTCT---CGGGAGCTCATGCTG 559  
QY 585 CAGGTGCTGCTCTCCAGTGCATCATCTGTACAGCTCATGCTCTTCTCTCGAGTAC 644  
DB 560 CAGATGCTGCTGCTGAGTGCATCATCTGTACAGCTCATGCTCTTCTCATGTTGATG 619  
QY 645 CGCGCGCGCGCGCTGCTCTCGACCAAGTTTCCCGACGCGCGCGCGCTCCATGCTC 704  
DB 620 CGCGCGCGCGGATGCTCATCACCGAGCATTCCCGGA---CACCGCGCGCAACATCGCC 676  
QY 705 TCTTTCGGGTGATTCGACGCTGCTCTGCTCGCAGGGGGAGCTGAGCTCGAGGCC 764  
DB 677 TCCATGCTGCTGACCCCGACGCTGCTGTGTGGAGCGGAGGAGCGCATCGAGACG 736  
QY 765 GAGCCCGACGGCTGCGCGCGCGCGCTCTCTCTCCCGCGGGGAGCGCGGGCGG 824  
DB 737 GAGACGGA-----GCTGAAGAGGACGCGGAGG 763

QY 825 GTGCGGTACCGTGGCGAAGTCCACAGCTCGCGTCCGAGCGCGCTGCTCGCACTC- 883  
DB 764 ATACAGTACCGTGGCGCGCTCCACAGCGTTCGCTCGGACATCTACTCCCGCGCTCC 823  
QY 884 --GACTCCAGACCATGACAGCCCGTGTGTGTCACCTCTTCGGGCTGTGAGATCTACTCG 941  
DB 824 ATGGGCTTCTCCAGCACACCGCGCGCGGAGCACTTCCACCAACCGGAGATCTACTCG 883  
QY 942 CTCAGTGTGTGCGGCAACCCACCCCGCGCGGTCCAGTTCACACCAACCGGACTTCTTC 1001  
DB 884 CTCAGTGTGTGCGGCAACCCAGCGCGGAGGTTCAAGCTTCAACCAACCGGACTTCTAC 943  
QY 1002 AACATGTCGGCGCCCGCCCAAGGAGGAGCGGAGGAGCGCGGGGAGCAGGAGAGGGC 1061  
DB 944 TCCATGTTGGCGCGCAGCTCCAACTTCGGCGCGGCGGACGCGTTCGCGCTCCGACCGGC 1003  
QY 1062 GC----- 1063  
DB 1004 GCCACGCGGCGCCGTCCTCACTACGAGGAGCAGCGGTCCAGGCCCAAGTACCGGCTCCG 1063  
QY 1064 -----ATGCGCGCGCGCGCGGAGGAC-ATCGCGCGACCGCGCGCGCTTCG- 1111  
DB 1064 GCGTCGAATGCGCGCGCCATGCGCGGCACACTACCGCGCGCGGAAACCGCGCGCTGCTCG 1123  
QY 1112 -----CGTCCGCGCAAGGAGGAGCTGAC 1139  
DB 1124 GCGCCCAAGGCGCCCAAGAGCGCGCCACGAAACGCGGCAAGGCGGAGGACCTCCAC 1183  
QY 1140 ATGCTGCTGTGAGCTTCCAGCGCTCCGCGCGTTCGAGCGCGCGCGCTGACGCTTTC 1199  
DB 1184 ATGTTGCTGTGAGCTTCCAGCGCGTTCGCGCGTTCGCGTTCG-----ACGCTTC 1228  
QY 1200 GGGCGCGCGCGCTGACCATGCGGAGCTCTCGCCAAAGGAGCCCGAGGCTTACGAGAG 1259  
DB 1229 GGGCGCGCGCGCAGACTACAAACGCGCGCGGAGTCAAGTCCCGCGCAAAATGGAT 1288  
QY 1260 TAGCGCGCGCAGCTACAGCAGCAGAGAGAGAGAGCGGCGCGCGCGGCAAGAGGCG 1319  
DB 1289 GAGCGAAGGACAGGAGGAGCTACGTTGAGCGGAGCATTCAGCTTCGGAACAGCGCG 1348  
QY 1320 G---GGCGCAGCTGTGAGGCTGGGTCCAACTCGACGCGCGAGCTGTACCCCAAGGAC 1376  
DB 1349 GTCATGACAGGAGACCGGAGGACGAGAGGCGGCGCGCGCGCGCGCGCGCGCGCGAC 1408  
QY 1377 GAGCGGAGGAGGCGCGCGGTGGCGATGCGCGCGCGCGAGCTGTATGACGCGGCTC 1436  
DB 1409 CCAGCAAGCCATGGCGCGCGCGAGCGGATGCGCGCGAGCGGTGATGACCGGCTTC 1468  
QY 1437 ATCTCATCATGTTGTGAGGAGCTGATCCGGAACCCCAACACCTTCTCCAGGCTCATC 1496  
DB 1469 ATCTCATCATGTTGTGAGGAGCTCATCCGCAACCCCAACACCTTCTCCAGGCTCATC 1528  
QY 1497 GCGGTGCTGTGCTGCTGCTTCTTACAGTGGGCGATCGAGATGCGAGCATCATCGCC 1556  
DB 1529 GCGCTCATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1588  
QY 1557 CGGTGATTTGATCTGTCGAGCGGCTCTCGGATGGCCATGTTTCAAGCTTACGCTG 1616  
DB 1589 AATATCATTCGATCTGTGCGAGCGGCGCTCGGATGGCCATGTTTCAAGTCTCGGCTG 1648  
QY 1617 TCCATGCGCTGAGCGGAGGATCATCGGTGCGGGAACAAGTGGCGGCGCATCGGATG 1676  
DB 1649 TCCATGCGCTGAGCGGCGCATCATCGGTGCGGGAACAAGTGGCGGCGCATCGGCGATG 1708  
QY 1677 GCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1736  
DB 1709 GCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1768  
QY 1737 GCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1796  
DB 1769 CGTGGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1828  
QY 1797 TTGCTGTTCCGCAAGGAGTACGCGGTTTCATCCCGACATCTCTGAGCACAG 1846



Db 1829 TTCGTCTTCCCAAGGAGTACACGCTGCACCTTAGCATTTCTCAGCACAGC 1878  
RESULT 6  
AAA94733  
ID AAA94733 standard; DNA; 2293 BP.  
XX AC AAA94733;  
XX DT  
XX 02-FEB-2001 (first entry)  
XX DE Wheat auxin transport protein clone wdk1c.pk008.g1 DNA sequence.  
XX KW Auxin transport protein; wheat; root development; gene mapping;  
XX KW plant breeding; herbicide; ss.  
XX OS Triticum aestivum.  
XX PN W0200068389-A2.  
XX XX  
XX PD 16-NOV-2000.  
XX PF 03-MAY-2000; 2000MO-US012061.  
XX XX  
XX PR 07-MAY-1999; 99US-0133040P.  
XX PA (DUPO) DU PONT DE NEMOURS & CO E I.  
XX PA (PION-) PIONEER HI-BRED INT INC.  
XX XX  
XX PI Orozco EM, Weng Z, Bruce WB, Cahoon RE, Tao Y;  
XX XX  
XX XX WPI; 2000-687647/67.  
XX DR P-PSDB; AAB26946.  
XX XX  
XX PT New nucleic acid sequences encoding new auxin transport proteins, useful  
XX PT for modulating root growth of plants and to screen for herbicides.  
XX XX  
XX PS Claim 2; Page 84; 94pp; English.  
XX CC Auxins are plant hormones that influence plant behaviour and development  
XX CC e.g. vascular tissue differentiation, apical development, tropic  
XX CC responses and organ (e.g. flower, leaf) development. The present  
XX CC invention relates to corn auxin transport protein coding sequences and  
XX CC proteins. The present sequence is one such coding sequence. This sequence  
XX CC may be used to modulate root development, e.g. to produce a more robust  
XX CC root system, alter root angle or redirect root growth. Also, the present  
XX CC sequence may be useful for gene mapping (e.g. for plant breeding) and to  
XX CC identify loss of function mutants. The protein encoded by the present  
XX CC sequence may be useful for raising specific antibodies, for the detection  
XX CC of auxin transport proteins and to design and/or identify specific  
XX CC inhibitors of auxin transport proteins, potentially useful as herbicides  
XX SQ Sequence 2293 BP; 462 A; 713 C; 683 G; 435 T; 0 U; 0 Other;  
Query Match 22.6%; Score 626.8; DB 3; Length 2293;  
Best Local Similarity 64.5%; Pred. No. 3.8e-91;  
Matches 1199; Conservative 0; Mismatches 467; Indels 192; Gaps 10;  
QY 164 CGCGGCCATGATCACCAGCGTGCACCTTACACGCTGCTGACGCGGCTGTCGCTGTA 223  
Db 67 CGTCCGGATGATCACCAGGGAAGACATCTACGACGTGCTGGCGGCTGTCGCTGTA 126  
QY 224 CGTGGCCATGACCTGGCGTACGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 283  
Db 127 CGTGGCCATGATCACCAGGGAAGACATCTACGACGTGCTGGCGGCTGTCGCTGTA 186  
QY 284 GTGCTCGGGGATCAACCGCTTCTGCGGCTGCTTTCGCGGCTGCTTCTCTCTTCCACTT 343  
Db 187 GTGCTCGGGGATCAACCGCTTCTGCGGCTGCTTTCGCGGCTGCTTCTCTCTTCCACTT 246  
QY 344 CATCTCCACCAAGACCCCTTCCGCTATGAACCTGCGCTTCTGCGGCTGCTGCGGCTGCA 403

Db 247 CATCTCCACCAACGACCCCTTAGCCATGAGCTACCGCTTCTTGGCCGCCGACGCTGCA 306  
QY 404 GAAGGTGGCCGCTCCTGCGCTGCTGGCGCTGGGCTCCCGCGGCTTCTCTTCCCGCGCGC 463  
Db 307 GAAGCTCGTCACTCGCGCGCTCGCGCTGCGGCAACAAGTGTCTCTCCCGCTACCGGTG 366  
QY 464 GCTCGG-----GCTCGACTGGAGCATCAGCTCTTCTC 496  
Db 367 CCGCGCGGACAGGAGCCCGGAGGCTCGTCTGAGACTGGACCATCAGCTCTTCTC 426  
QY 497 CCTCTCACGCTCCCAACACGCTCGTATGGGATCCCGCTGCTGCGAGGATGTACGG 556  
Db 427 CCGCGGACGCTGCCCAACACGCTGCTGATGGGATCCCGCTGCTGCGCGCTATGACGG 486  
QY 557 CGGCTCGTGGCGGACGCTCATGCTCAGGTCGCTCTCTCAGTGCATCTCTGTA 616  
Db 487 CGACTTCTCG---GGTCTCGTCTGTCAGATGCTGCTGTCAGAGCTCATCTCTGTA 543  
QY 617 CACGCTCATGCTCTTCTCTTTCAGATACCGCGCGCGCGCTCTCTCTCGACCAAGTT 676  
Db 544 CACGCTCATGCTCTTCTCTTTCAGATACCGCGCGCGCGCTCATCTCTCGACGAGTT 603  
QY 677 CCGGACGCGCGCGCGCTCATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 736  
Db 604 CCGCGCGGACGCTCGGCGCGGACATCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 663  
QY 737 CGCAGGCGGACGCTCGAGCTCGAGGCGCGGACCGCGCTCGCGCGCGCGCGCGCGCT 796  
Db 564 CAACGCGCGGAGCGCTGCAACCGCGCGGAGTGGCGCGGACGCGCGCTCCACGT 723  
QY 797 CTCCTCCGCGCGGAGCGCGCGGCGGCTCGCGCTCACCGCGCGCGCGCGCGCGCTCCCG 856  
Db 724 CGTCACTCGCGCGGCTCGCGCTCGCGCTCGCGCTCGCGCGCGCGCGCGCGCTCCCG 783  
QY 857 GCGCTCGAGCGCGCGCTGCTCGCACTCGCACTCCAGACCATGCGCGCGCGCGCGCGCT 916  
Db 784 GATCTACGCTGGCGG-----TTCACGCGCATGACGCGCGCGCGCTCAA 828  
QY 917 CCTCTCCGCGTGGAGATCTACTCGTGCAGTCTCGTGCAGAACCCACCGCGCGCGGCT 976  
Db 829 CTTCAAGCGCGTGGAGATCTACTCGCTGCAGAGCTCGCGGAGCCACGCGCGGAGTC 888  
QY 977 CAGCTTCAACCGCGCGCTCTCTCAACATCTCGCGG----- 1014  
Db 889 CAGCTTCAACCGCGCGCTCTCTCAACATCTCTCAACGCGAGCAAGTGTGCTAGTCCAA 948  
QY 1015 -----CGCGCGCCAAAGGAGCGGAGCGGCGGCGG-----GACGAGGAGAA 1057  
Db 949 GGGCCAGCGCGCGCTCGCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1008  
QY 1058 GGGCGCATGCGGCGCGCGCGGAGGACACTCGCGCGCGCGCGCGCGCTGCGCGT-- 1115  
Db 1009 GGGCAACAAGTTCAAGGGCGGCGAGGCGGCTGCGCTTACCGCGCGCGCGCGCGGAT 1068  
QY 1116 -----CGGCGCCAAAGGAGGAGGA 1132  
Db 1069 GATGATCGCGCGCGCACGGAAGAGAGCTTGGGGGTTCCAACTCAAACTCGAACAGGA 1128  
QY 1133 CTTGCAATGCTGTGTGAGCTCCAGGCTCGCGCTCGCGGCTCGCGGCGCGCGCGCTGCA 1192  
Db 1129 GCTGCAATGCTGTGTGAGCTCCAGCGCTCGCGCTCGCGGCTCGGAGGCCAACCTCGCAA 1188  
QY 1193 ---CGTCTTGGGCGCGCGCGCTGACCAATGCGGAGCTCTCGCCAAAGAGCCAGGC 1249  
Db 1189 CGCGCTCAACCAACGCGCGCTCCACCGACTTTCGCGCGCGCGCGCGCGCGCGCGCG 1248  
QY 1250 CTACGACAGTACGGGCGCGACGACTACAGCAGG-----GACGAGAA 1294  
Db 1249 ACGAGAGCGGCGGACACCCAGAGCGGTGAGCGGCGGTGACGCGGTGATGAGAGGA 1308  
QY 1295 CGGAGCGGCGGCGGAGCAAGGCGG-----GGCGGCGGTGTC 1333  
Db 1309 CGCAGGAGCGGCGGCTGAGGTGAGATCGAGGACGGCATGATGAGAGCGCGCGGAC 1368



Db 776 TGGCGGCTCGAGCTTCGGCGCTCCGAGCTCTACTCGTCAATCGTCGGGGCCCAA 835  
Qy 1042 CGGGGACAGGAGAGGCGCATCGCGCGCGGGCGGAGACACTCGCGCAGCGCG 1101  
Db 836 CCCCAGGAGCTCCACTTCGAGGAGCACTCGGACGCGCGGAAACACCGCAACA 895  
Qy 1102 AGGCGGTGCCGTCGCGGCAAGAGAGAGCACTGCAATGCTGCTGTGAGCTCCAGC 1161  
Db 896 CCACGGGGCACTCAACACGATGCGCAAGAGCTCCACATGTTCTGTGTGAGCTCGAGC 955  
Qy 1162 CTTCCGCGCTGTCGAGCGCGCGCGCTGCTCTTC---GGCGCGCGCGCTGACC 1218  
Db 956 CGTCTCCCGTCTCAGAAGTACGCGGCTGCTGTGTTGAGTGGCGGCGCGCGCGG 1015  
Qy 1219 ATGCGGAGCTCTCGCAAGAGAGCCCA----- 1246  
Db 1016 CTCTGAGCTCGCGCGCAAGGAATCCACATGTCATCCCGCGCACTCGCGCAGAAC 1075  
Qy 1247 -----GGCCTACGAGC 1257  
Db 1076 AGGCTCAGGCAAGAGCAGAGGATCGGCGGAGTGGGTGGCGCGCGCGCG 1135  
Qy 1258 AGTACGGCGCGACGACTACAGCAGCAGGAGCAAGAACCGGAGCGCGCGCGCAAG 1317  
Db 1136 GAGAGAACTTACGTTCCGAGGCGGCAAGCGTGGAGCGCGCGGAGGAGTAGACGAG 1195  
Qy 1318 GCGGGC-----GACGCTGTCAAGCTGGGTCAACTCGAGCGCGAGCTGTACC 1368  
Db 1196 AGCGCGCTTGCCTGACGCGGCTGACGAAGATGGGTGCGAGCTCGACGCGGAGCTGCACC 1255  
Qy 1369 CCAAG-----GACGCGCGGAGGAGCGCGCGCGG 1401  
Db 1256 CGAAGTCTGCGAGCTCGAGGACGAAACCGCGCGCGCGCGCGCGCGCGAGT 1315  
Qy 1402 TGGCGATGCGCGCGGAGGCTGATGACGCGGCTCATCTCATGATGTTGAGGAGAGC 1461  
Db 1316 ACCAAATGCGCGCGGAGGCTGATGACAGCGCTCATCTCATTAATGTTGTTGGCGCAAGC 1375  
Qy 1462 TGATCGGAGCCCAACACTACTTCCAGCTCATCGGCTGCTGCTGCTGCTGCTGCT 1521  
Db 1376 TCATCCGCAACCCCAACACTTACTTCAGGCTCTCGGCTCGCTGCTGCTGCTGCT 1435  
Qy 1522 ACAGTGGGCGATCGAGATGCCAGGATCATGCGCGCTGATTCGATTCCTGTGAGCG 1581  
Db 1436 TCGGTGGCAGCTTCCTACGCGAGCAATCGTCGAGAAGTCCATCTCCATTCCTCGGAGC 1495  
Qy 1582 CGGCTCTCGGATGGCATGTTTCAAGCTAGGCTGTTTCAAGGCTGCGAGCGAGGATCA 1641  
Db 1496 CAGGCTTGGGATGGCGATGTTTAGGCTGGGATGTTTCAAGGCTGCGAGCGAGGATCA 1555  
Qy 1642 TCGGTCGGGGAACAAGCTGCGCGCCATCGCGATGGGCGTTCGGTTCGTCGAGGCGCG 1701  
Db 1556 TCGGTCGGCAAAATCAGCGCGCTGCTCTCCATGCGCGTTCGCTTCCTCGGCGCGCTG 1615  
Qy 1702 CGGTATGCGCGGCTCCATCGCGTGGTCTGCGCGCGTTCCTCCATCGCCA 1761  
Db 1616 CCGTCATGGCGCGCGCTCAATCGCATCGGATCGCGGAGCGCTTCCTGCGAGTGGCCA 1675  
Qy 1762 TCGTCAGGCTGCTCTGCTTCAAGGATCGTGGCTTCTGTTTTCGCAAGGAGTACGGCG 1821  
Db 1676 TTGTTTCAAGCGGCTTACACACAGGATGTTGCTTTTGTGTTTTCGCAAGGATCAATG 1735  
Qy 1822 TTGATCCGAGATCTGAGCAGAGCG 1847  
Db 1736 TCCACCGCGCATCTGAGCAGAGCG 1761

RESULT 8  
ADA69396  
ID ADA69396 standard; DNA; 1845 BP.  
XX  
AC  
ADA69396;  
XX

DT 20-NOV-2003 (first entry)  
XX Rice gene, SEQ ID 2719.  
XX Plant; bacterial infection; fungal infection; viral infection; rice;  
XX gene; ds.  
XX Oryza sativa.  
XX WO2003000898-A1.  
XX 03-JAN-2003.  
XX 22-JUN-2001; 2001WO-1B001105.  
XX 22-JUN-2001; 2001WO-1B001105.  
XX (SYGN) SYNGENTA PARTICIPATIONS AG.  
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
XX Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;  
XX WPI; 2003-175290/17.  
XX Identifying at least one gene involved in plant resistance or response to  
XX pathogenic infection for conferring resistance or tolerance to a plant to  
XX bacterial, fungal or viral infection by determining or detecting plant  
XX gene expression.  
XX Claim 6; SEQ ID NO 2719; 899pp; English.  
XX The present invention relates to a method (M1) for identifying genes  
XX involved in plant resistance or response to pathogenic infection. M1  
XX comprises identifying a gene whose expression is significantly altered in  
XX the incompatible interaction of plant gene expression relative to  
XX expression of the gene in an uninfected plant, in a mutant plant that  
XX does not express a gene associated with response to pathogenic infection,  
XX or in a corresponding incompatible or compatible interaction. (M1) is  
XX useful for conferring resistance to resistance or tolerance to a plant to  
XX bacterial, fungal or viral infection. The present sequence was used to  
XX illustrate the invention.  
XX Sequence 1845 BP; 315 A; 628 C; 557 G; 345 T; 0 U; 0 Other;

Query Match 20.3%; Score 563; DB 7; Length 1845;  
Best Local Similarity 63.5%; Pred. No. 5.8e-8;  
Matches 1158; Conservative 0; Mismatches 430; Indels 237; Gaps 10;  
Qy 211 TGGTGGCGCTCTAGTGGCCATGACGCTGGCTACGCTCCGCTGGTGGCGCATCT 270  
Db 2 TGGTGGCGCTCTAGTGGCCATGACGCTGGCTACGCTCCGCTGGTGGCGCATCT 61  
Qy 271 TCACGCGGACCAAGTCTCGGGATCAACCGCTTCGTGGCGCTTCGCGGTGCGGTCC 330  
Db 62 TCACGCGGACCAAGTCTCGGGATCAACCGCTTCGTGGCGCTTCGCGGTGCGGTCC 121  
Qy 331 TCTCTTCCCTTCTATCTCCACCAACGACCCCTTCGCGCATGAACCTTCCTGGCGG 390  
Db 122 TGTGTTCCATCTTCTTCCACCAACCGGTACGATGACCTCGGTTTCATCGCG 181  
Qy 391 CCGACACGCTCGAAGAGTGGCCCTCTCGCGCTGCTGGCGCTGGCTTCGCGCGCTCT 450  
Db 182 CCGACACGCTCGAAGAGTGGCCCTCTCGCGCTGCTGGCGCTGGCGCGCTTCAGC- 240  
Qy 451 CCTCCCGCGCGCTCGGCTCGACTGGAGCATCAGCTCTTCTCCTCTCCACGCTCC 510  
Db 241 -----CGCGCGGGAGCCTCGAGTGGACCATCAGCTCTTCTCCTCTCCAGCTGC 292  
Qy 511 CCAACACGCTCGTATGGGATCCCGCTGCTGCGAGGATGACGCGCGCTGCTGGCGCG 570  
Db 293 CCAACACGCTCGTATGGGATCCCGCTGCTGCGAGGATGACGCGCGAGTTCT---CG 349  
Qy 571 GCACGCTCATGGTCAGGTGCTGCTGCTCCAGTGCATCATCTGTTACACGCTCATGCTCT 630

[illegible]

Db	1394	TCATCCGCAACCCGAAACACCTACTCCAGCGCTCATCTGGCTCATCTGGTCCCTCGTCTGCT	1455
Qy	1522	ACAGGTGGGGCATCGAGATGCCAGCGCATCATGCCCGGTGAGATTTTCGATCCTCTFCGAGC	1581
Db	1454	TCAGGTGGAACCTTCGAGATGCGCGCCATCGTCTGAAATCCATCTCGATCCTCTGCGAGC	1513
Qy	1582	CGGGTCTCGGGATGGCCATGTTTCAGCCTAGGCTGTTTCATGGCGCTGCAGCCGAGGATCA	1641
Db	1514	CGGGGTTCGCGCATGGCCATGTTTCAGTCTCGGTCTGTTTCATGGCGCTGCAGCCGACATCA	1573
Qy	1642	TCGCTGTCGGGAAACAAGCTGGCGGCCCATCGCGATGGCGCTTCGGTTCGTCGAGGCCCG	1701
Db	1574	TCGCTGTCGGGAAACAAGGTGGCGCACTAGCCATGCGGTTCTGCGGTTCTTGCCGGGCCG	1633
Qy	1702	CGGTATGCGCGCGCGCTTCATCGCGTTCGGTTCGGCGCGCTTCCTCCACATCGCCA	1761
Db	1634	CCGTGATGGCGCGCGCTTCCTTCGCGCTCGGACTTCGCTGGCAGCGTCTGCAGTTCGCCA	1693
Qy	1762	TCGTCACAGGTGCTCTCGCTTCAGGGGATCGTGGCGTTTCGTTTCGCCAAGGAGTACGCG	1821
Db	1694	TTGTCCAGGCAGTCTGCCCGAGGCAATGTCCTTCGTTTCGCCAAGGAGTACAGCG	1753
Qy	1822	TTCAATCCGACATCTCGAGCACAGC	1846
Db	1754	TGCACCTTAGCATTTCTCAGCACAGC	1778
RESULT 9			
AAA94724			
Id	AAA94724 standard; DNA; 2162 BP.		
XX	AAA94724;		
XX	02-FEB-2001 (first entry)		
DE	Rice auxin transport protein clone rslln.pk003.n3 DNA sequence.		
XX	Auxin transport protein; rice; root development; gene mapping;		
KW	plant breeding; herbicide; ss.		
XX	Oryza sativa.		
XX	W0200068389-A2.		
PN	16-NOV-2000.		
FD	03-MAY-2000; 2000WO-US012061.		
XX	07-MAY-1999; 99US-0133040P.		
PR	(DUPO ) DU PONT DE NEMOURS & CO E I.		
XX	(PION-) PIONEER HI-BRED INT INC.		
FA	Orozco EM, Weng Z, Bruce WB, Cahoon RE, Tao Y;		
PI	WPI; 2000-687647/67.		
XX	P-PSDB; AAB26937.		
DR	New nucleic acid sequences encoding new auxin transport proteins, useful		
PT	for modulating root growth of plants and to screen for herbicides.		
XX	Claim 2; Page 65-66; 94pp; English.		
PS	Auxins are plant hormones that influence plant behaviour and development		
XX	e.g. vascular tissue differentiation, apical development, tropic		
CC	responses and organ (e.g. flower, leaf) development. The present		
CC	invention relates to corn auxin transport protein coding sequences and		
CC	proteins. The present sequence is one such coding sequence. This sequence		
CC	may be used to modulate root development, e.g. to produce a more robust		
CC	root system, alter root angle or redirect root growth. Also, the present		
CC	sequence may be useful for gene mapping (e.g. for plant breeding) and to		
CC	identify loss of function mutants. The protein encoded by the present		
CC	sequence may be useful for raising specific antibodies, for the detection		

XX New nucleic acid sequences encoding new auxin transport proteins, useful  
PT for modulating root growth of plants and to screen for herbicides.

PS Claim 2: Page 65-66; 94pp; English.

Auxins are plant hormones that influence plant behaviour and development e.g. vascular tissue differentiation, apical development, tropic responses and organ (e.g. flower, leaf) development. The present invention relates to corn auxin transport protein coding sequences and proteins. The present sequence is one such coding sequence. This sequence may be used to modulate root development, e.g. to produce a more robust root system, alter root angle or redirect root growth. Also, the present sequence may be useful for gene mapping (e.g. for plant breeding) and to identify loss of function mutants. The protein encoded by the present sequence may be useful for raising specific antibodies, for the detection

CC of auxin transport proteins and to design and/or identify specific  
XX inhibitors of auxin transport proteins, potentially useful as herbicides  
SQ Sequence 2162 BP; 428 A; 694 C; 636 G; 404 T; 0 U; 0 Other;

```
Query Match          19.4%; Score 536.2; DB 3; Length 2162;
Best Local Similarity 59.4%; P-Val 1.1e-76;
Matches 1021; Conservative 0; Mismatches 658; Indels 40; Gaps 5;

QY 163 CCGCGCCATGATCAACGGCTGGAGCTCTACACAGTGTGACGGCGGTGGTGGCGCTGT 222
DB 1123 AGAGGAGGACGCTGCACATGCTGTGAGCTCCAGCGCTCGCCGTTGTCGAGCGCG 1182
QY 91 CTGAGGAGATGATATCCGGGACGACTTCTACAGGTGATGGGGGTGGTGGCGCTGT 150
DB 1045 ATGCCAAGGAGCTCCACATGTTCTGTGGAGCTCGAGCGCTCTCCGTTCTCAGAAGTCA 1104
QY 223 AGTGGCCATGAGCGTGGGTGACGGTCCGTCGGTGGTGGGCGATCTTACCGCGGACC 282
DB 1183 CCGCGGTGACGCTCTTC---GGCGCGCGCGCGCTGACCATGCGGACGCTCTCGCAAAG 1239
QY 151 AGTGGCGATGTTCTTGGGCTACGGGTGCGGTGGTGGGCGCATCTTACCGCGGACC 210
DB 1105 GCGGCTTGCCTGTGTTCAAGTGGCGCGCGCGCGCGCGCTCTCGACGCTCGGCGCAAAG 1164
QY 283 AGTGTCTCCGGGATCAACCGCTTCTGGGCGCTCTTCCGCGTGGCGCTCTCTCTCTTCACT 342
DB 1240 GAGCCACAGGCTTACGAC-----GAGTACGGGCGCG 1269
QY 211 AGTGTCTCCGGCATCAACCGTTCGTGCGCATCTTTCGGCGTGGCGCTCTGTCTCTCACT 270
DB 1165 AATTCACATGTTGTCATCCCGCGGACCTCGCGCAGAACACGGCTCAGGCAAGAGCAGC 1224
QY 343 TCATCTCCACCAACGACCCCTTTCGCCATGAACCTCGCTTCTTGGCGCGCGACAGCTGC 402
DB 1270 ACGACTACAGCAGCAGGACGAGAACGGGAGCGCGCGCGGAGCAAGGGCGGCGCGACGC 1329
QY 271 TCATCTCCACCAACGACCCGTAACGCTTCTCCGCTTCTCCAGCTGCGGACAGCTGC 330
DB 1225 AGAGTACCGCGCAGTGGCATTTGGGTGGCGCGCGCGCGCGGAGAGAACTTCAGCTTCGAG 1284
QY 403 AGAAGTGGCCGCTCTCGGCGCTCTGGCGCTTGGCGCTCGCGGCGCTCTCTCCCGCGCG 462
DB 1330 TGTGGAAGCTGGGGTCCAACTCGACCGCGCAGCTGTACCCCAAGGACGACGCGGAGGGGA 1389
QY 331 AGAAGTGTCTCTTGGGCGCTCGCGCGTGGCGCTTGGCGCTTGGCGCGCGGACAGCTGC 390
DB 1285 GCGGCAAGAGCGTGGAGCGGCGCGCGGAGCAGTAGACGAGGAGCGCGCTTGCCTGACGGGC 1344
QY 463 CGCTCGGCTCGATGGAGCATCAGCTCTTCTCCCTCTCCAGCTTCCCGGCGCGCTCGCTCAGTGG 582
DB 1390 GGGCGCGCGCGTGGCGATGCGCGCGCGGAGCGGTGATGACGCGGCTCATCTCTCATGCG 1449
QY 451 TCATGGGATCCCGCTGCTGATGCGCATGTAACG---GCCATCTCCGCGCTCGCTCATGG 507
DB 1345 TGACGAAGATGGGGTGGAGCTCGACCGCGGAGCTGCAACCGGAGGTCTGCGACGTGACG 1404
QY 583 TCAGGTCTCGTCTCCAGTGCATCATCTGTACACGCTCATGTCTTCTCTTCTTCGAGT 642
DB 1450 TGTGGAAGAGCTGATCCGGAACCCCAACACCTACTTCCAGCTCTATCGGCGCTCTGCTGGT 1509
QY 508 TCAGATCTGTGCTTCCAGTGCATCATCTGTACAGCTGATGTCTTCTCTTCTTCGAGT 567
DB 1405 GACCGAACGCGCGCGCGCGCGCGCGGCGGCGGAGTACCAATGCGCGCGCGGAGCG 1464
QY 643 ACGGCGCGCGCGCGCTGCTCTGACGAGTTCGCGAGCGCGCGCGCGCGCGCGCTGCTGCTG 702
DB 1510 CCTGTGCTCTCTACAGTGGGGCATCGAGATGCGACGATCATTCGCGCGCGGTGCAATTCGA 1569
QY 568 TCGCGCGCGCGGATGCTGATCGCGACCACTTCCCGA---CACGCGCGCGCTCCATCG 624
DB 1465 TGATGACACGCGCTCATCTCATTAATGGTGTGGCGCAAGCTCATTCGCAACCCCAACACTT 1524
QY 703 TCTCTCTCGCTCGACTCCGACGCTGCTCTCGCTCGCGAGGGGACGCTCGAGCTCGGAGG 762
DB 1570 TCTGTGCGGACGCGGTCTCGGATGGCCATGTTTCAGCTT-AGGCGCTGTTCATGGCGCTG 1628
QY 625 TGTCTCTGACGCTGACCGCGAGCTGCTGCTGAGGGGCGGCCACGCGGAGACGAGG 684
DB 1525 ACTCGAGCTCTCGGCTCGCGCTCGCTGCTCTCGCTTCGGAATGTTTCATGGCGCTG 1584
QY 763 CCGAGCCCGACGCGCTGCGCGCGCGCGCTGCTCTCTTCTCGCGCGCGGAGACGCGCGCG 822
DB 1629 CAGCGGAGGATCATCGCTGCGGGAACAAGCTGGCGGCCATCGCGATGGGCGTCCGGTTC 1688
QY 685 CCGAGGTGGCGGACGCGCGCGCTGCACTCACCGTGCAGCTCGCGCTCGCTCGCTCGCGCTG 744
DB 1585 CAGCCAGCATCATCGCTGTGGCAATCAGCGCGCGCTCTCTCCATGGCGCGTCCGCTTC 1644
QY 823 GGTGCGCGCTACCGTGGCGCAAGTCCACGACTCGGCTCGAGGCGCGGCTCGCACT 882
DB 1689 GTGCGGCGCGCGCGCTGATGGCGCGCGCTCCATCGCGTGGTCTGCGCGCGCTCTC 1748
QY 745 GGTGCTGTGTGTCAGCGCGCGCGCTGCACTGACGGGAGCGAGATCTACTCGCTTA 804
DB 883 CGCACTCCAGACCAATGACGCCCGCTGTGTCGAACTCTTCCGCGGTGGAGATCTACTCGC 942
QY 805 GCTCTGCGGGAACCAACCCCGCGGGGTCTCAACTTCAACCAACGCGACTTCTTCGCA 864
DB 943 TGCAGTCTGCGGCAACCCCAACCGCGGGGTCTGAGTTCAGTTCACCAACGCGACTTCTTCA 1002
QY 865 TGTGCGCGCGCGCGCGCGCGCGCGCGCGCTGCTCTCTTCTCGCGCGCGGCTCGAGCTTCGCG 924
DB 1003 ACATCGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1062
QY 925 CCTTCGAGCTTTACTGCTGCAATCGTTCGCGGGGCCCAACCCCGAGGAGTCCAACTTCG 984
DB 1063 CATGCGCGCGCGCGCGCGGAGACACTCGCCCGAGCGCGCGCGCTGCGCGTTCGCGCGCA 1122
QY 985 ACGAGCACTCGGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1044
```

## RESULT 10

AAA94720

ID AAA94720 standard; DNA; 504 BP.

XX

AC AAA94720;

XX

DT 02-FEB-2001 (first entry)

XX

DE Corn auxin transport protein EST clone p0119.cmtn124r DNA sequence.

XX

KM Auxin transport protein; corn; expressed sequence tag; EST;

XX

KM root development; gene mapping; plant breeding; herbicide; ss.

XX

OS Zea mays.

XX

PN W0200068389-A2.

XX

PD 16-NOV-2000.

XX

PP 03-MAY-2000; 2000WO-US012061.

[illegible]

AAA94719	AAA94719 standard; cDNA; 1426 BP.
ID	AAA94719 standard; cDNA; 1426 BP.
XX	AAA94719;
XX	02-FEB-2001 (first entry)
XX	Corn auxin transport protein clone p0094.cssh17r DNA sequence.
XX	Auxin transport protein; corn; root development; gene mapping;
XX	plant breeding; herbicide; ss.
XX	Zea mays.
OS	
XX	Key Location/Qualifiers
PH	CDS 320..1426
FT	/*tag= a
FT	/product= "Corn auxin transport protein clone
FT	p0094.cssh17r"
FT	/note= "No stop codon given"
XX	
XX	W0200068389-A2,
XX	16-NOV-2000.
XX	03-MAY-2000; 2000WO-US012061.
XX	07-MAY-1999; 99US-0133040P.
XX	(DUPO ) DU PONT DE NEMOURS & CO E I.
PA	(FION-) PIONEER HI-BRED INT INC.
XX	Orozco EM, Weng Z, Bruce WB, Cahoon RE, Tao Y;
XX	WPI; 2000-687647/67.
DR	P-PSDB; AAB26932.
XX	
XX	New nucleic acid sequences encoding new auxin transport proteins, useful
XX	for modulating root growth of plants and to screen for herbicides.
XX	Claim 2; Page 56; 94pp; English.
XX	
XX	Auxins are plant hormones that influence plant behaviour and development
CC	e.g. vascular tissue differentiation, apical development, tropic
CC	responses and organ (e.g. flower, leaf) development. The present sequence
CC	is a cDNA insert comprising of clone p0094.cssh17r. This sequence
CC	encodes a corn auxin transport protein. This sequence may be used to
CC	modulate root development, e.g. to produce a more robust root system,
CC	alter root angle or redirect root growth. Also, the present sequence may
CC	be useful for gene mapping (e.g. for plant breeding) and to identify loss
CC	of function mutants. The protein encoded by the present sequence may be
CC	useful for raising specific antibodies, for the detection of auxin
CC	transport proteins and to design and/or identify specific inhibitors of
CC	auxin transport proteins, potentially useful as herbicides
XX	
XX	Sequence 1426 BP; 241 A; 528 C; 415 G; 242 T; 0 U; 0 Other;
XX	
XX	Query March 15.4%; Score 425.6; DB 3; Length 1426;
XX	Best Local Similarity 72.1%; Pred. No. 5.2e-59;
XX	Matches 646; Conservative 0; Mismatches 199; Indels 51; Gaps 5;
Qy	171 ATGATCACCGCGTGGACCTCTACACGCTCTACCGCGGTGTGCGCTGTACGTGGCC 230
Db	320 ATGATTACGGGGACGGACTTCTACACGCTCATACCGCGGTGTGCGCTGTACGTGGCG 379
Qy	231 ATGACGTCGTGCGTACGGCTCCGTCGCTGGTGGCGATCTTTCACGGCGGACCAAGTGTCC 290
Db	380 ATGATCTCGGCTACGGGTGCGGTGGTGGCGCATCTTCTCGCGGACCAAGTGTCTCC 439
Qy	291 GGGATCAACCGCTTCGTGGCGCTCTTTCGCGGTGCGCGTCTCTCTCTCATCTCATCTCC 350
Db	440 GGGATCAACCGCTTCGTGGCGCTCTTTCGCGGTGCGCGTCTGTCTCTCATCTCATCTCC 499

RESULT 11



```
QY 351 ACCAAGACCCCTTCCCATGAACCTGCGCTTCTGGCGCGGACACGCTGCGAAGGTG 410
    |||||
Db 500 ACCAACAACCCCTTACACCTGAACCTGCGCTTCTGGCGCGGACACGCTGCGAAG--- 556
    |||||
QY 411 GCGGTCTCTGCGCTGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 470
    |||||
Db 557 -----CTAGTGTGTGGCCATGTCTACCGGTGGAGCCACTCAGCGCGGGGCGGCG 610
    |||||
QY 471 CTGACTGGAGCATCACGCTCTTCTCCCTCTCCAGCTCTCCCAACACGCTCTCATATGGGC 530
    |||||
Db 611 CTGGAGTGGACCATCACGCTCTTCTCCCTCTCCAGCTCTCCCAACACGCTCTCATATGGGC 670
    |||||
QY 531 ATCCGCTGCTGGAGGAGCATGTACGGCGGTGTGGCGCGGACGCTCATATGGGTCCAGGTC 590
    |||||
Db 671 ATCCGCTGCTCAAGGAGCATGTACGGCGGACTTCT---CGGAGGCTCATGGTGGAGATC 727
    |||||
QY 591 GTGCTCTCCAGTGCATCATGTGTACAGCTCATGTCTTCTCTTGGAGTACCGCGCC 650
    |||||
Db 728 GTGCTGCTCCAGTGCATCATGTGTACAGCTCATGTCTTCTTATGTTGAGTACCGCGCC 787
    |||||
QY 651 GCGCGCGCGCTGCTCTCGACACAGTTTCCCGGAGCGCGCGCTCATATGCTTCTCTTC 710
    |||||
Db 788 GCGCGGATGCTCATACCGGAGAGTTTCCCGGA---CAAGCGCGGGGCCATCGCTTCCATC 844
    |||||
QY 711 CGGTCGACTCCGAGCTGCTCTGCTCGCCGCGGGGAGCTGAGTCCGAGCCGAGGCC 770
    |||||
Db 845 GTGCTCGACCGGAGCTGCTCTCCCTCGAGCGCGGAGGAGCGCCATCGAGACGGAGGCC 904
    |||||
QY 771 GAGCGGCTGCGCGCGCGCGCGCTCTCTCCCGCGGGGAGCGCGCGGGTGGGC 830
    |||||
Db 905 GA-----GGTCAAGGAGGAGCGGAGGATACAC 931
    |||||
QY 831 GTACACGTGGGAGTCCAGCAGCTGCGCTCGGA---GGCGCGTGTCTGCACTCGCAC 887
    |||||
Db 932 GTACCGTGGCGGCTTCCACGCTCGCTCGGATCTACTCGCGCGCTTCCATGGGC 991
    |||||
QY 888 TCCACAGCATGACGCGCGCTGTGTCCAACTCTCGGGGTGGAGATCTACTCGCTGGAG 947
    |||||
Db 992 TTCTCAGCAGCACGCGCGCGCCAGCAACCTGACCAACGCGGAGATCTACTCGCTGGAG 1051
    |||||
QY 948 TCGTCGGCAACCCACCGCGCGGCTCGAGTTCAACACGCGGACTTCTTCAACATC 1007
    |||||
Db 1052 TCGTCGGCAACCCACCGCGCGGCTCGAGTTCAACACGAGACTTCTTACTCCATG 1111
    |||||
QY 1008 GTCGGCGCGCGCGCCCAAGGAGGAGCGGAGGAGCGCGGGGAGGAGGAGGCGCG 1063
    |||||
Db 1112 GTCGGCGGAGCTCCAACTTCGCGCGGCGGAGCGGCTTGGCATCGGACCGGCGC 1167
    |||||

RESULT 12
AAC43229
ID AAC43229 standard; DNA; 1851 BP.
XX
XX AAC43229;
AC
XX 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 38501.
DE
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
XX Arabidopsis thaliana.
OS
XX EP1033405-A2.
PN
XX 06-SEP-2000.
PD
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
PR
XX 05-MAR-1999; 99US-0123180P.
PR
```

```
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136332P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140931P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
```



```
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149923P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.

PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 12.5%; Score 347.2; DB 3; Length 1851;
Best Local Similarity 54.9%; Pred. No. 1.8e-46;
Matches 989; Conservative 0; Mismatches 673; Indels 138; Gaps 10;

Qy 171 ATGATCACCAGCGCTGGACCTCTTACACGCTGCTGACGGCGGTGCTGCGGTGACGTGGCC 230
    |||||
Db 1 ATGATTACGTGGCAGACTTGTACACCGCTCTACCGCGGTGCTGCGGTGACCTTTACGTAGCT 60
    |||||

Qy 231 ATGACGCTGGCGTACGGCTCCGCTCGCTGGTGGCGCATCTTCACGCCGACACGAGTGTCTCC 290
    |||||
Db 61 ATGATTCTCGCTACCGATCCGTACAGTGGTGAAGATATTCTCACCAGACGAGTCTCC 120
    |||||

Qy 291 GGGATCAACGGCTTGTGGCGGCTTTTGGCGGTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCC 350
    |||||
Db 121 GGCATCAACGGCTTGTGGCGGCTATCTTGGCGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCC 180
    |||||

Qy 351 ACCAAGACCCCTTGGCCATGAACCTGGCTTCCCTGGCGCGGACACACGCTGCAGAGGTG 410
    |||||
Db 181 ACCAAGATCCTTACCGCATGATTTCCGTTGTCGCGCGGACACACGCTTCAAAAATC 240
    |||||

Qy 411 GCGCTCTCGCGTGTGGCGGTCCCGCGGCTCTCTCTCCCGCGGCTCTCTCTCCCGCGGCTCGGG 470
    |||||
Db 241 ATCATGCTGTCTTACTTGTCTCT-----ATGGGCTAACCTAACCAAGACGCT-AGC 291
    |||||

Qy 471 CTCGACTGGAGCATCAGCTCTTCTCCCTCTCCAGCTCCCAACACGCTCTCTCATGGGC 530
    |||||
Db 292 TTGGAGTGGATGATCAAACTTCTCTCAGCATCTCTCCAAACACTTCTTTCATGGGG 351
    |||||

Qy 531 ATCCCGCTCTCGGAGGACATGACGGCGGTGCTGCGCGGACGCTCTATGFTCCAGGTG 590
    |||||
Db 352 ATCCCTCTGTTGATGCCATGTACGG---AACCTACGAGGTCTCTAATGGTCCAAAGTC 408
    |||||

Qy 591 GTGCTCTCCAGTGCATCATCTGTGATACGCTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCC 650
    |||||
Db 409 GTTGTCTCTCAGTGTATCATTTGGTACACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 468
    |||||

Qy 651 GCGCGCGCGCTGCTCTCTCGACCAAGTTCCCGGAGCGCGCGCGCTCTCTCTCTCTCTCTCTCTCT 710
    |||||
Db 469 GCTAAGCTTCTTATTATGAGGAGTTCCTCGGA---GACTGTGCGTCTATTGTTTCGTTT 525
    |||||
```







QY 351 ACCAACGACCCCTTCGCGATGAACCTCGCTTCCTGGCGCGGACACGCTCGAGAAGTG 410  
Db 181 TCRAACAATCCCTTACGAATGAATCTCCGATTCATCGCAGCGGATACACTCCAAACTA 240  
QY 411 GCCGCTCGCGCTCTCGCGCTCGCGCTCCGCGCGCTCTCTCTCCCGCGCGCTCGG 470  
Db 241 ATCATGCTCACTCTCTCT-----AATCATCTGGGCTAACTTCCTCTCGCTCGGAGT 291  
QY 471 CTCGACTCGAGCATACCGCTCTTCCTTCCTTCCTCCAGCTCCCAACAGCGCTCGTATGGC 530  
Db 292 CTGGAATGAGCATCACAATCTTCCTCTCTTCAACACTTCCCAACTCTCGTTATGGGA 351  
QY 531 ATCCGCTCTCTCGAGGATGACGCGGCTCGTCGCGCGGACGCTCATGCTCCAGGTC 590  
Db 352 ATACCTCTCTTCATCGCCATGACGAGAAATATT---CCGGCTCACTCATGCTTCAGATC 408  
QY 591 GTCGCTCTCCAGTGCATCATCTGGTACAGCTCATGCTCTTCCTTCCTTCGAGTACGCGCC 650  
Db 409 GTTGCTCTTCAGTGATTAATCTGGTACACGCTTCTCTCTCTTCCTTCGAAATACAGAGA 468  
QY 651 GCGCGCGGCTCTCTCTCGACAGTTCGCGAGCGCGCGCGCTCCATGCTCTCTCTTC 710  
Db 469 GCTAAGATCTTCATCATGGAACAGTTTCAGA---GACTGGTCTTCGATGATCTGTTTC 525  
QY 711 CCGCTCGACTCCGACGTCTCTCGCTCGCGAGGGGAGCTCGAGCTCGAGCGCGAGCCC 770  
Db 526 AAAGTTGATCGGATGTTGTTCTTTTAGATGACATGATTTTCTTGAACACTGATGCTCAA 585  
QY 771 GACGCGCTCGCGCGCGCGCGCTCTCTCTCCCGCGCGGAGCGCGCGGCTGGC 830  
Db 586 ATCGGTGACGATGTTAAGCTTCATGTTAGGTTGAGAAATCAACGCTTCTC-GGAGATC 644  
QY 831 GTACCGTGGCAATGCCACAGCTCGCGCTCCGAGCGCGGCTCGCACTCG-----884  
Db 645 TTTTACGCTGGTGGTGGTACTAAATGATGACTCTCGTCTGCTAACTCTCACCGGAGTGA 704  
QY 885 -----CACTCCAGACCATGACG-----903  
Db 705 GATTTATAGCTTAACACTACTCCAGAGAGATCTAACTTCAATCTCTGATTTTATCT 764  
QY 904 -----CCGCTGTGTCCAACCTCTCCGCGCTGGAGATCTACTCGCT 943  
Db 765 TATGATGGGTTTCCCGGTGGCGGCTTTTGAACCTTTGGTCCGCGGATATGTACTCTGT 824  
QY 944 GAGTCTGTCGCGCAACCCACCCCGCGGCTCGAGCTTCAACACGCGGACTTCTTCAA 1003  
Db 825 TCAGTCTCTCGAGGACCTACTCTCTAGACCTTTCGAAATTTGAGGA-GAGTTGTCTATGG 883  
QY 1004 CATCTGCGCGCGCGCGCAAGGGAGCGAGGAGCGCGGGGACGAGGAGAGGGGCG 1063  
Db 884 CTCTCTCGCAAGATTCGGGTATTACCGGGAGGAGCTCCCGGGTCTGTACCCAGCTCTTA 943  
QY 1064 ATCGCGCGCGCGCGGAGGACACTCGCGCGAGCGCGCGCGCTCGCGCGGCG---1120  
Db 944 ATCCGGAGTTTTCTACCGGTAAATAAACCGGTAGTAAAGCTCCAAAGAGAAATCATCATC 1003  
QY 1121 -----CAAGGAGGAGCTTCACATGCTCTCTGGAGCTCCA 1158  
Db 1004 ATGTAGGAAAAATCAATAGTAATGATGCTAAGAGCTTCATATGTTTGTGTTGGGATCCA 1063  
QY 1159 GCGCTCTCGCCGCTCGGAGCGCGCGCTGTCACGT-----CTTCGGCG 1203  
Db 1064 ACGGATCACCCGTTTCGGACCGAGCTGGTCTTCAAGTTGATAATGAGCAATGAACAAG 1123  
QY 1204 CCGCGCGGCTGACCATGCGAGCTCTCTCGCCAAAGAGGCCAGGCTTACGACGAGTACG 1263  
Db 1124 TCGGAAATCCGATCAAGGCGGTGCAAAAGAGATTGCAATGTTGATCTCTGATCATATCTC 1183  
QY 1264 GCGCGGACCATCTACAGCAGCAGAGAGAAACGAGGAGCGCGCGGACGAGGCGGCG 1323  
Db 1184 AAAATGCTGAACAAAGCTGTCTCCGATGAACGCGGACTATGCGCGGGAAGAGAGTCGG 1243  
QY 1324 CGAG-----CTCTCGAAGCTGGGCTCCAACTCGAGCGGCGGAGC 1362

Db 1244 AGAGGTAAGGAAGTGCTTAACGGAATACACAGCTTCGGTGTAACTCCACAGCAGAGC 1303  
QY 1363 TGTACCCCAAGGACGAGCGGAGGAGGCGCGCGGCTGG-----CGATGCCGCGG 1416  
Db 1304 TAAACCCCTAAGGAAGCTATAGAAAACGGGTGAACCTGTACCGGTAAACATATAGCCACCAG 1363  
QY 1417 CGAGCGTGTATGACGCGGCTCATCTCATCATGTGTGGAGGAAGCTGATCCGGAACCCCA 1476  
Db 1364 CGAGTGTGATGACTCGGCTGATATTGATTAATGGTGGAGGAACTCATAGAAGAACCCAA 1423  
QY 1477 ACACCTACTCCAGCCTCATCGGCGTGTGTGCTCTGTGCTCTCTTACAGGTGGGGCATCG 1536  
Db 1424 ACACCTACTCTAGTCTCATTTGGGCTCTTGTGGGCTCTTGTGCTTTCAGGTGGGATGTGG 1483  
QY 1537 AGATGCGAGCATCATCGCGGCTCGATTTTCGATCTCTGCGACGCGGCTCTCGGATGG 1596  
Db 1484 CAATGCTTAAATATTCAACATCAATCTCAATCTTCTGATGCTGTGCTTGTGATGG 1543  
QY 1597 CCATGTTCAAGCTTAGGCTCTTTCATGCGCTGTGAGCGGAGGATCATCGGCTGGGGAACA 1656  
Db 1544 CAATGTTCAAGTTGGGCTTTCATGGCATTTGCAACCGAAATTAATTTGCTTGGGAAAT 1603  
QY 1657 AGTGGGCGCATCGCATGGGCTCGGTTTCTGTCGAGGCCCCGCGGCTCATGGCGCGG 1716  
Db 1604 CGACGCGCATTTTTCGATGGCGGTGAGATTTCTTACTGACAGCGGTAATGGCCGTGG 1663  
QY 1717 CTCTCATCGCGCTCGGCTCTGCGCGGCTCTCTCTCCACATCGCCATCGTCCAGGCTGCTC 1776  
Db 1664 CAGCAATGGCTATTGGATTACGTGGAGACCTATTGGTGTGGCCATTGTTCAAGCTGCAT 1723  
QY 1777 TGCCTCAGGGATCGTGGCTTGGTTCGCGGAGGAGTACGCGGTTTCATCCGACATCC 1836  
Db 1724 TGCTTCAAGGATCGTGGCGGCTTGTGTTTTCGAAAAGAGTATAATGTTTCATCCGCAATCT 1783  
QY 1837 TGAGCACAGCG 1847  
Db 1784 TGAGTACAGG 1794

Search completed: March 3, 2004, 01:25:04  
Job time : 1062 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run On: March 3, 2004, 01:10:58 ; Search time 6898 Seconds  
(without alignments)  
11987.297 Million cell updates/sec

Title: US-10-030-884-13

Perfect score: 2769

Sequence: 1 ccacgcgtccgctgagccct.....aaaaaaaaaaaaaaaaaaag 2769

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:

1: em\_estba.\*

2: em\_esthum.\*

3: em\_estin.\*

4: em\_estmu.\*

5: em\_estov.\*

6: em\_estpl.\*

7: em\_estro.\*

8: em\_hic.\*

9: gb\_est1.\*

10: gb\_est2.\*

11: gb\_hic.\*

12: gb\_est3.\*

13: gb\_est4.\*

14: gb\_est5.\*

15: em\_estfun.\*

16: em\_estom.\*

17: em\_gss\_hum.\*

18: em\_gss\_inv.\*

19: em\_gss\_pln.\*

20: em\_gss\_vrt.\*

21: em\_gss\_fun.\*

22: em\_gss\_mam.\*

23: em\_gss\_mus.\*

24: em\_gss\_pro.\*

25: em\_gss\_rod.\*

26: em\_gss\_phg.\*

27: em\_gss\_vrl.\*

28: gb\_gss1.\*

29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2108	76.1	2737	11	AY110494
2	877.8	31.7	893	29	CG376782
3	840	30.3	840	29	CG358070
4	726	26.2	726	29	CG358058

C	5	711	25.7	730	29	CG376770	CG376770
	6	706	25.5	717	14	CF244269	CF244269
	7	639	23.1	835	29	CG722061	CG722061
	8	619.4	22.4	658	13	BU098540	BU098540
C	9	599	21.6	958	28	CG371169	CG371169
	10	576.8	20.8	610	13	BU080190	BU080190
	11	574.4	20.7	630	14	CA830783	CA830783
	12	569	20.5	627	13	BU037419	BU037419
	13	568.4	20.5	794	29	CG456385	CG456385
	14	561	20.3	776	29	CG317974	CG317974
	15	559.8	20.2	604	14	CA831110	CA831110
	16	527.4	19.0	815	28	BZ644423	BZ644423
	17	518.4	18.7	552	13	BU037594	BU037594
	18	510	18.4	510	14	CF244718	CF244718
	19	477.6	17.2	497	13	BU036965	BU036965
	20	431	15.6	469	14	CD001628	CD001628
C	21	429	15.5	857	29	CG057348	CG057348
	22	429	15.5	891	29	CG057347	CG057347
	23	425.6	15.4	1426	11	AY106004	AY106004
	24	413.8	14.9	1107	14	CK209475	CK209475
C	25	399	14.4	831	28	BZ644430	BZ644430
	26	396.2	14.3	903	14	CF243672	CF243672
	27	395.4	14.3	1105	14	CK208792	CK208792
	28	385.4	13.9	1109	14	CK208849	CK208849
	29	383.4	13.8	428	13	BU500094	BU500094
	30	382.8	13.8	386	28	BH801495	BH801495
	31	380.4	13.7	1221	11	AY110024	AY110024
C	32	380.2	13.7	924	29	CG293496	CG293496
	33	378.6	13.7	722	14	CD934785	CD934785
	34	376.4	13.6	619	14	CA3236508	CA3236508
C	35	364.6	13.2	533	28	BH873074	BH873074
	36	359.8	13.0	696	14	CA184766	CA184766
	37	356.4	12.9	739	14	CB646582	CB646582
	38	356.4	12.9	819	14	CB648277	CB648277
	39	350.4	12.7	703	29	CG637427	CG637427
	40	349.8	12.6	899	14	CA227062	CA227062
	41	346.2	12.5	585	13	CA144854	CA144854
	42	344.6	12.4	731	14	CB668851	CB668851
	43	342.2	12.4	631	13	BQ295136	BQ295136
	44	339.6	12.3	621	13	CA099711	CA099711
	45	338.8	12.2	694	14	CA192422	CA192422

#### ALIGNMENTS

RESULT 1	AY110494	2737 bp	linear	HTC 17-OCT-2002
LOCUS	AY110494	2737 bp	linear	HTC 17-OCT-2002
DEFINITION	Zea mays CL464_1 mRNA sequence.			
ACCESSION	AY110494			
VERSION	AY110494.1	GI:21214903		
KEYWORDS	HTC.			
SOURCE	Zea mays			
ORGANISM	Zea mays			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.			
REFERENCE	1 (bases 1 to 2737)			
AUTHORS	Rainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whittitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.			
TITLE	Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes			
JOURNAL	Unpublished (2002)			
REFERENCE	2 (bases 1 to 2737)			
AUTHORS	Coe, E.H.			
TITLE	Direct Submission			
JOURNAL	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA			
COMMENT	If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the			





```
QY 1861 CATCGCATGTTTTCATCTGCTGATAGTAAACGGGAAAAAAGCAGACGATCGATG 1920
Db 1861 CATCGCATGTTTTCATCTGCTGATAGTAAACGGGAAAAAAGCAGACGATCGATG 1920
QY 1921 ACAGCGCATGAATTCATCTGATGATTAATGATGATGATGATGATGATGATGATG 1980
Db 1921 ACAGCGCATGAATTCATCTGATGATTAATGATGATGATGATGATGATGATGATG 1980
QY 1981 AAAGAACCACTAATGACATGATCTAGGACAGATCAGATGATGATGATGATGATG 2040
Db 1981 AAAGAACCACTAATGACATGATCTAGGACAGATCAGATGATGATGATGATGATG 2040
QY 2041 TCCTGACAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2100
Db 2041 TCCTGACAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2100
QY 2101 TACAACACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2160
Db 2101 TACAACACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2160
QY 2161 CATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220
Db 2161 CATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220
QY 2221 TCGCTTCTTCAGCTGCGGAGCGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280
Db 2221 TCGCTTCTTCAGCTGCGGAGCGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280
QY 2281 TGATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340
Db 2281 TGATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340
QY 2341 CGGGACTGCGGGAGTCAAGACAGACCCCTCGAAACAAAAGTTTCTTTTGGCTTCTG 2400
Db 2341 CGGGACTGCGGGAGTCAAGACAGACCCCTCGAAACAAAAGTTTCTTTTGGCTTCTG 2400
QY 2401 GTCAAGAACAAAGTTTGGCTTTGGCTTTGGCTTTGGCTTTGGCTTTGGCTTTGG 2460
Db 2401 GTCAAGAACAAAGTTTGGCTTTGGCTTTGGCTTTGGCTTTGGCTTTGGCTTTGG 2460
QY 2461 ATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2520
Db 2461 NNNNTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2520
QY 2521 CCAGAGGCAATGGAATCAACAAAGCTTCTAGCGCGGCTATCATATCAAGGAACACA 2580
Db 2521 CCAGAGGCAATGGAATCAACAAAGCTTCTAGCGCGGCTATCATATCAAGGAACACA 2580
QY 2581 TGCAGATACAAACGAGTCTAGTGCAGAAATGCTTCTTTTCTTTTCTTTTCTTTTCT 2640
Db 2581 TGCAGATACAAACGAGTCTAGTGCAGAAATGCTTCTTTTCTTTTCTTTTCTTTTCT 2640
QY 2641 GGGTTTCTAGCTGATTAAGGATTCAGATGATGATGATGATGATGATGATGATGATG 2700
Db 2641 GGGTTTCTAGCTGATTAAGGATTCAGATGATGATGATGATGATGATGATGATGATG 2700
QY 2701 ACAATTTCTGGCTTTTGAAGAAATTCCTCTCGTT 2737
Db 2701 ACAATTTCTGGCTTTTGAAGAAATTCCTCTCGTT 2737

RESULT 2
CG376782
LOCUS CG376782 893 bp DNA linear GSS 26-AUG-2003
DEFINITION OGCN22TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMA0776C20,
genomic survey sequence.
ACCESSION CG376782
VERSION CG376782.1 GI:34294049
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

```
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES
source
1..893
/organism="Zea mays"
/mol_type="genomic DNA"
/strains="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0776C20"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 31.7%; Score 877.8; DB 29; Length 893;
Best Local Similarity 99.7%; Pred No. 9,6e-80;
Matches 890; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 103 ACAAAATTGGCAATGCTGTCGCGGACCCCTCTCTCCCTCCCGCGTCCCGGACGAT 162
Db 1 ACAAAATTGGCAATGCTGTCGCGGACCCCTCTCTCCCTCCCGCGTCCCGGACGAT 60
QY 163 CGCGGCGCATGATACCGGCTGGACCTTACACGCTGTACGCGGTGTGCGCGTGT 222
Db 61 CGCGGCGCATGATACCGGCTGGACCTTACACGCTGTGTACGCGGTGTGCGCGTGT 120
QY 223 AGCTGGCCATGACGCTGGCGTACGCGTCCGTCGCTGTGTGTGCGCATCTTACGCGGAC 282
Db 121 AGCTGGCCATGACGCTGGCGTACGCGTCCGTCGCTGTGTGTGCGCATCTTACGCGGAC 180
QY 283 AGTGTCTCGGATCAACCGCTTCTGTCGCGCTTCTGCGCTGCGCTCTCTCTTCCACT 342
Db 181 AGTGTCTCGGATCAACCGCTTCTGTCGCGCTTCTGCGCTGCGCTCTCTCTTCCACT 240
QY 343 TCATCTCCACCAACGACCCCTTCGCGCATGAACCTGCGCTTCTGCGCGCGCGACGCTGC 402
Db 241 TCATCTCCACCAACGACCCCTTCGCGCATGAACCTGCGCTTCTGCGCGCGCGACGCTGC 300
QY 403 AGAAGTGGCGCTCTCTGCGCTGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 462
Db 301 AGAAGTGGCGCTCTCTGCGCTGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 360
QY 463 CGCTCGGCTCGACGTGAGCATCAAGCTTCTCTCTCTCCACGCTCCCGGACGCTCG 522
Db 361 CGCTCGGCTCGACGTGAGCATCAAGCTTCTCTCTCTCCACGCTCCCGGACGCTCG 420
QY 523 TCATGGGCATCCCGTGTGTGCGGATGATGACGCGCTGCTGCGCGCGCGCTCATGG 582
Db 421 TCATGGGCATCCCGTGTGTGCGGATGATGACGCGCTGCTGCGCGCGCGCTCATGG 480
QY 583 TCCAGTGTCTCTCTCCAGTGCATCTGTGTACAGCTCATGTCTTCTCTTTCGAGT 642
Db 481 TCCAGTGTCTCTCTCCAGTGCATCTGTGTACAGCTCATGTCTTCTCTTTCGAGT 540
QY 643 ACCGCGCGCGCGCGCTGCTGCTCTCGACAGTTCCCGACGCGCGCGCGCTCCATCG 702
Db 541 ACCGCGCGCGCGCGCTGCTGCTCTCGACAGTTCCCGACGCGCGCGCGCTCCATCG 600
```

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 893)

Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,  
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,  
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.

Consortium for Maize Genomics

Unpublished (2002)

Other GSSs: OG3CN22TH

Contact: Cathy Whitelaw

7912 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TF

Class: sheared ends.

Location/Qualifiers

1..893

/organism="Zea mays"

/mol\_type="genomic DNA"

/strains="B73"

/db\_xref="taxon:4577"

/clone="ZMMBMA0776C20"

/clone\_lib="ZM\_0.7\_1.5\_KB"

/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb

methylation filtered genomic DNA library"

```

QY 703 TCTCCTTCGGCGTGCAGTCCGACGTCGTCTGCTCGCTCGCCAGGGGGAGCGTCGAGCTCGAGG 762
Db 601 TCTCCTTCGGCGTGCAGTCCGACGTCGTCTGCTCGCTCGCCAGGGGGAGCGTCGAGCTCGAGG 660
QY 763 CGAGCCCGAGCGGCGTCCGCGCGCGGCGCGGCGGCTCTCTCCGCGGGGAGCGCCGCGG 822
Db 661 CGAGCCCGAGCGGCGTCCGCGCGCGGCGCGGCGGCTCTCTCCGCGGGGAGCGCCGCGG 720
QY 823 GGGTGGCGGTCAACCGTGGCGAAGTCACACAGCTCGCGCTCGGAGCGCGCGTGTGCGCACT 882
Db 721 GGGTGGCGGTCAACCGTGGCGAAGTCACACAGCTCGCGCTCGGAGCGCGCGTGTGCGCACT 780
QY 883 CGCACTCCAGACCAATGACGCCCGGTGTGTCAACCTCTCCGCGGTGGAGATCTACTCGC 942
Db 781 CGCACTCCAGACCAATGACGCCCGGTGTGTCAACCTCTCCGCGGTGGAGATCTACTCGC 840
QY 943 TCCAGTCTGTCGCGCGCAACCCCA-CCCCGGGCGGGGTCCAGCTTCAACACGCGGA 994
Db 841 TCCAGTCTGTCGCGCGCAACCCCAACCCCGCGGGTCCAGCTTCCACACGCGGA 893

RESULT 3
CG358070/c
LOCUS
DEFINITION
CG358070
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 840)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other-GSSs: OGI.BK16TH
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.

FEATURES
source
1..840
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBma0724C07"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 30.3%; Score 840; DB 29; Length 840;
Best Local Similarity 100.0%; Pred. No. 6.5e-76;
Matches 840; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 388 CGCGCCACACGCTGCAGAGGTGGCGTCTCTCGCGCTGTGCGCTGCGCCTCCGCGGCC 447
Db 840 CGCGCCACACGCTGCAGAGGTGGCGTCTCTCGCGCTGTGCGCTGCGCCTCCGCGGCC 781
QY 448 TCTCCTCCCGCGCGGCTCGGCTCGACTGAGATCAAGCTTCTTCCCTCTCCAGC 507
Db 780 TCTCCTCCCGCGCGGCTCGGCTCGACTGAGATCAAGCTTCTTCCCTCTCCAGC 721

```

```

QY 508 TCCCAACACGCTCGTCAATGGGCAATCCCGCTGCTCGAGGCATGTACGGCGGCTGCTGG 567
Db 720 TCCCAACACGCTCGTCAATGGGCAATCCCGCTGCTCGAGGCATGTACGGCGGCTGCTGG 661
QY 568 CGGGCACGCTCATGGTCCAGGTCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 627
Db 660 CGGGCACGCTCATGGTCCAGGTCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 601
QY 628 TCTTCTCTTTCGAGTACCGCGCGCGCGCGCGCTGCTCTCTCTCTCTCTCTCTCTCTCT 687
Db 600 TCTTCTCTTTCGAGTACCGCGCGCGCGCGCGCTGCTCTCTCTCTCTCTCTCTCTCTCT 541
QY 688 CGGGCACGCTCATGGTCCAGGTCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 747
Db 540 CGGGCACGCTCATGGTCCAGGTCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 481
QY 748 ACCTCGAGTCTCGAGCGCGCGCGCGCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 807
Db 480 ACCTCGAGTCTCGAGCGCGCGCGCGCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 421
QY 808 CGGGCACGCGCGCGCGCGCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 867
Db 420 CGGGCACGCGCGCGCGCGCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 361
QY 868 CGCGTGTCTCGCACTCGCACTCCAGACCATGCGCGCGCGCGCTGTGTCCAACTCTCCGG 927
Db 360 CGCGTGTCTCGCACTCGCACTCCAGACCATGCGCGCGCGCGCTGTGTCCAACTCTCCGG 301
QY 928 TGGAGATCTACTCGCTGCGAGTGTGCGCAACCCACCGCGCGCGCTGCTCTCTCTCTCT 987
Db 300 TGGAGATCTACTCGCTGCGAGTGTGCGCAACCCACCGCGCGCGCTGCTCTCTCTCTCT 241
QY 988 ACGCCCACTTCTTCAACATCTGTCGGCGCGCGCGCGCAAGGGAGCGGAGGCGGGGG 1047
Db 240 ACGCCCACTTCTTCAACATCTGTCGGCGCGCGCGCGCAAGGGAGCGGAGGCGGGGG 181
QY 1048 ACAGAGAGAGGCGCGCATGCGCGCGCGCGCGCGAGAGACACTGCGCGCGAGCGCGCGCG 1107
Db 180 ACAGAGAGAGGCGCGCATGCGCGCGCGCGCGCGAGAGACACTGCGCGCGAGCGCGCGCG 121
QY 1108 TCGCGCTGCGCGCGCAAGAGAGAGGACCTGCATGCTCTCTCTCTCTCTCTCTCTCTCT 1167
Db 120 TCGCGCTGCGCGCGCAAGAGAGAGGACCTGCATGCTCTCTCTCTCTCTCTCTCTCTCT 61
QY 1168 CCGTGTCCGAGCGCGCGCGCGTGCAGCTTTCGCGCGCGCGCGCGCTGACCATGCCGAGC 1227
Db 60 CCGTGTCCGAGCGCGCGCGCGTGCAGCTTTCGCGCGCGCGCGCGCTGACCATGCCGAGC 1

RESULT 4
CG358058
LOCUS
DEFINITION
CG358058
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 726)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other-GSSs: OGI.BK16TV
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843

```

RESULT 5		CG376770		730 bp		DNA		linear		GSS 26-AUG-2003	
CG376770/c		OG3CN22TH ZM_0.7.1.5 KB		Zea mays		genomic clone		ZMMBma0776C20,			
LOCUS		genomic survey sequence.									
DEFINITION		ACCESSION		VERSION		KEYWORDS		SOURCE			
		CG376770		CG376770.1		GI:34294037		GSS.			
		Zea mays		Zea mays		Zea mays		Zea mays			
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.		1 (bases 1 to 730)							
		Whitelaw, C.A., Quackenbush, J., Van Aken, S., Uterback, T., Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.									
		Consortium for Maize Genomics									
		Unpublished (2002)									
		Other GSSs: OG3CN22TV									
		Contact: Cathy Whitelaw									
		TIGR									
		9712 Medical Center Drive, Rockville, MD 20850, USA									
		Tel: 301-838-5843									
		Fax: 301-838-0208									
		Email: whitelaw@tigr.org									
		Seq primer: TR									
		Class: Sheared ends									
		Location/Qualifiers									
		1. .730									
		/organism="Zea mays"									
		/mol_type="genomic DNA"									
		/strain="B73"									
		/db_xref="taxon:4577"									
		/clone="ZMMBma0776C20"									
		/clone_lib="ZM_0.7.1.5 KB"									
		/note="Vector: pBGSK-; Site 1: HincII; 0.7-1.5 kb									
		methylation filtered genomic DNA library"									
		ORIGIN									
		Query Match		25.7%;		Score 711;		DB 29;		Length 730;	
		Best Local Similarity		100.0%;		Pred. No. 7.3e-63;					
		Matches 711;		Conservative 0;		Mismatches 0;		Indels 0;		Gaps 0;	
		QY		530		CATCCCGTGTGCGAGCATGTATCGGCGCTGCTCGCGCGGACGTCATGTCAGGT		589			
		Db		730		CATCCCGTGTGCGAGCATGTATCGGCGCTGCTCGCGCGGACGTCATGTCAGGT		671			
		QY		590		CGTCTGCTCCAGTGCATCATCTGTACAGCTCATGCTCTTCTCTTCAGTACCGCGC		649			
		Db		670		CGTCTGCTCCAGTGCATCATCTGTACAGCTCATGCTCTTCTCTTCAGTACCGCGC		611			
		QY		650		CGCGCGCGCTGCTGCTCGACAGTTCCTCCGACGGCGCGCGCGCTCCATGCTCTCTT		709			
		Db		610		CGCGCGCGCTGCTGCTCGACAGTTCCTCCGACGGCGCGCGCGCTCCATGCTCTCTT		551			
		QY		710		CGCGCTCCACTCCGAGCTGCTGCTCGCGCGGAGGAGTTCAGAGTCCGAGGCGGAGCC		769			
		Db		550		CGCGCTCCACTCCGAGCTGCTGCTCGCGCGGAGGAGTTCAGAGTCCGAGGCGGAGCC		491			
		QY		770		CGACGCGCTGCGCGCGCGCGCTGCTCTCTCCCGCGCGGAGCGCGCGCGGCTGCG		829			
		Db		490		CGACGCGCTGCGCGCGCGCGCTGCTCTCTCCCGCGCGGAGCGCGCGCGGCTGCG		431			
		QY		830		CGTCAACGCTGCGAAGTCCACAGCTCGCGTCCGAGGCGCGCTGCTGCGACTCGCATC		889			
		Db		430		CGTCAACGCTGCGAAGTCCACAGCTCGCGTCCGAGGCGCGCTGCTGCGACTCGCATC		371			
		QY		890		CCAGACCTGCGAGCCCGCTGCTCCAACTCTCCGGGTGAGATCTACTCGCTGCGATC		949			
		Db		370		CCAGACCTGCGAGCCCGCTGCTCCAACTCTCCGGGTGAGATCTACTCGCTGCGATC		311			
		QY		950		GTCGCGCAACCCCAACCCCGCGCGGTCCAGCTTCAACCAACCGCGCTTCTTCAACATCGT		1009			





Db	125	CGAGTACGGCGCGACGACGTACAGCAGCAGGAAACGGGAGCGCGCGCGGACAA	184
Qy	1316	GGCGGGCCCGACGCTGTGCGAGCTGGGGTCCAACTCGACCGCGCAGCTGTACCCCAAGGA	1375
Db	185	GGCGGGCCCGACGCTGTGCGAGCTGGGGTCCAACTCGACCGCGCAGCTGTACCCCAAGGA	244
Qy	1376	CGACGGCGAGGGAGGGCGCGCGGTGGCGATCCCGCGCGAGCTGTGATCGCGGCT	1435
Db	245	CGACGGCGAGGGGAGGACGGCGC--GGCGATCCCGCGCGAGCTGTGATCGCGGCT	301
Qy	1436	CATCCTCATCATGTGTGGAGGAAGCTGATCCGGAACCCCAACCTACTCCAGCCTCAT	1495
Db	302	CATCCTCATCATGTGTGGAGGAAGCTGATCCGGAACCCCAACCTACTCCAGCCTCAT	361
Qy	1496	CGGGGTGCTGTGTCCTGTCTCTTACAGTGGGGCATCGAGATGCGGGGATCATGCG	1555
Db	362	CGGGGTGCTGTGTCCTGTCTCTTACAGTGGGGCATCGAGATGCGGGGATCATGCG	421
Qy	1556	CGGGTCGATTTTCGATCTCTGTGCGACGGGGTCTCGGGATGGCCATGTTTCAGCCTAGGCGCT	1615
Db	422	CGGGTCGATTTTCGATCTCTGTGCGACGGGGTCTCGGGATGGCCATGTTTCAGCCTAGGCGCT	481
Qy	1616	GTTTCATGGCGGTGACGCCGAGGATCATGCGTGTGGGGAACGAGCTGGGGGCATCGGAT	1675
Db	482	GTTTCATGGCGGTGACGCCGAGGATCATGCGTGTGGGGAACGAGCTGGGGGCATCGGAT	541
Qy	1676	GGGGTCCGGTTCGTGCGAGGCCCGGGGTCAATGGCGCGCGCTCCATCGCGGTGGGTCT	1735
Db	542	GGGGTCCGGTTCGTGCGAGGCCCGGGGTCAATGGCGCGCGCTCCATCGCGGTGGGTCT	601
Qy	1736	GCGGGCGTCTCTCTCAATCGGCATGTCAGGCTGTCTGCTCAGGGGATCGT	1792
Db	602	GCGGGCGTCTCTCTCAATCGGCATGTCAGGCTGTCTGCTCAGGGGATCGT	658

RESULT 9	CC371169/c	CC371169	LOCUS	DEFINITION	958 bp	DNA	linear	GSS:19-MAY-2003
		CC371169	PURK16TD	ZM 0.6 1.0 KB	Zea mays	genomic	clone ZMMBtra523D07,	
				genomic survey		sequence.		

VERSION	CC371169.1	GI:30844785
KEYWORDS	GSS.	
SOURCE	Zea mays	
ORGANISM	Zea mays	
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.	

REFERENCE	1. (bases 1 to 958)
AUTHORS	Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and Bennettzen, J.
TITLE	Maize Genomics Consortium
JOURNAL	Unpublished (2003)
COMMENT	Other GSSs: PUMSK16TB Contact: Cathy Whitelaw TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-838-5843 Fax: 301-838-0208 Email: whitelaw@tigr.org Seq primer: TF Class: sheared ends.

FEATURES	source
Location/Qualifiers	1..958
	/organism="zebra mays"
	/mol_type="genomic DNA"
	/strain="B73"
	/db_xref="taxon:4577"
	/clone="ZMMB1523D07"
	/clone_lib="ZM 0.6 1.0 KB"
	/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
	Col selected genomic DNA library"

Query Match 21.6%; Score 599; DB 28; Length 958;  
 Best local similarity 88.8%; Pred. No. 1.3e-51;  
 Matches 706; Conservative 0; Mismatches 5; Indels 84; Gaps 2;  
 1608 CTAGGCTGTTCATGGCGCTCGACGCGAGGATCATCGCTGCGGGAACAAGCTGGCGGCC 1667  
 Db 882 CAGGCTGTTCATGGCGCTCGACGCGAGGATCATCGCTGCGGGAACAAGCTGGCGGCC 823  
 1668 ATGCGGATGGGCGTCCGGTTCGTGCGAGGCGCCGCGGTCTATGCGCGCGCTCCATCGCC 1727  
 Db 822 ATGCGGATGGGCGTCCGGTTCGTGCGAGGCGCCGCGGTCTATGCGCGCGCTCCATCGCC 763  
 1728 GTGCGTCTGCGCGCGCTCTCTCCACATCGCCATCG- 1764  
 Db 762 GTGCGTCTGCGCGCGCTCTCTCCACATCGCCATCGTCCAGGTACGCGCCCTGTGGCTT 703  
 1765 ----- 1764  
 702 GAGGTAAAGCGGTGAGTGTGATGGAAACAGGTAGGACCAACATTCGTTCCCTTCTTCT 643  
 1765 TCCAGGCTGTCTGCTCCTCAGGGGATCGTCCGTTTCGTGTTTCGCAAGGAGTACGGGTTTC 1824  
 Db 642 TCCAGGCTGTCTGCTCCTCAGGGGATCGTCCGTTTCGTGTTTCGCAAGGAGTACGGGTTTC 583  
 1825 ATCCGACATCTCGACACAGCGGTATGTTCCCAATAACATCGCATGGTTTTCATCACTTGCC 1884  
 Db 582 ATCCGACATCTCGACACAGCGGTATGTTCCCAATAACATCGCATGGTTTTCATCACTTGCC 523  
 1885 ATAGTTAAAGCGGAAAAAAGCAGAGCAATCGATGACGACGCACTGAATTCATATGAT 1944  
 Db 522 ATAGTTAAAGCGGAAAAAAGCAGAGCAATCGATGACGACGCACTGAATTCATATGAT 463  
 1945 TCATTACTATGATGGTGTTCATGTCAGTGCAGTCAAGAAACACCTAATAAGCACTGAT 2004  
 Db 462 TCATTACTATGATGGTGTTCATGTCAGTGCAGTCAAGAAACACCTAATAAGCACTGAT 403  
 2005 CTAGGACAGCATCAGCATGATGATGCTGTTTTCCTCTGCAAACTCGCATTCCTTACT 2064  
 Db 402 CTAGGACAGCATCAGCATGATGATGCTGTTTTCCTCTGCAAACTCGCATTCCTTACT 343  
 2065 ACAGGTGCGCTTCACATCCATCCAGATGATCATCAACACTACTGATGCATC-TTT 2123  
 Db 342 ACAGGTGCGCTTCACATCCATCCAGATGATCATCAACACTACTGATGCATCCTTTT 283  
 2124 TTTTTCATTCGTGACGCTGATCTTCGGGATGTCATCGCTCTGCCCATCACTCTGG 2183  
 Db 282 TTTTTCATTCGTGACGCTGATCTTCGGGATGTCATCGCTCTGCCCATCACTCTGG 223  
 2184 TCCTACTACTTCCTCGGCGTGTGACCTCTCTCGCTCGCTTCTTCAGCGTGGGGAAGG 2243  
 Db 222 TCCTACTACTTCCTCGGCGTGTGACCTCTCTCGCTCGCTTCTTCAGCGTGGGGAAGG 163  
 2244 CGCCATCTGTGTGTATTCGCGCCCCACATGAAATTTCTGATGAGCATTAGCATTTACTAT 2303  
 Db 162 CGCCATCTGTGTGTATTCGCGCCCCACATGAAATTTCTGATGAGCATTAGCATTTACTAT 103  
 2304 TAGCTTAGCGAAGAA 2318  
 Db 102 TAGCTTAGGTAGGGA 88  
 RESULT 10  
 BU080190  
 LOCUS  
 DEFINITION BU080190 610 bp mRNA linear EST 27-AUG-2002  
 945150G03.y1.946 - tassal primordium prepared by Schmidt lab Zea  
 mays cDNA, mRNA sequence.  
 ACCESSION BU080190  
 VERSION BU080190.1 GI:22521379  
 KEYWORDS EST.  
 SOURCE Zea mays  
 ORGANISM Zea mays

RESULT 10	BU080190	610 bp	mrna	linear	EST 27-AUG-2002
LOCUS	BU080190				
DEFINITION	945150G03.y1 946 - tassell primordium prepared by schmidt lab Zea mays CDNA, mRNA sequence.				
ACCESSION	BU080190				
VERSION	BU080190.1				
KEYWORDS	GI:22521379				
SOURCE	EST.				
ORGANISM	Zea mays				
	Zea mays				



Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.	
1 (bases 1 to 610)	
Walbot,V.	
Maize ESTs from various cDNA libraries sequenced at Stanford University	
Unpublished (1999)	
Contact: Walbot V	
Department of Biological Sciences	
Stanford University	
855 California Ave, Palo Alto, CA 94304, USA	
Tel: 650 723 2227	
Fax: 650 725 8221	
Email: walbot@stanford.edu	
Plate: 946150 row: G column: 03.	
Location/Qualifiers	
1. 610	
/organism="Zea mays"	
/mol_type="mRNA"	
/cultivar="OH43"	
/db_xref="taxon:4577"	
/tissue_type="tassels"	
/dev_stage="just after the transition from vegetative to inflorescence development"	
/lab_host="XL0LR"	
/clone_lib="946 - tassel primordium prepared by Schmidt lab"	
/note="Organ: tassels; Vector: HybrizAP; Site_1: EcoRI; Site_2: XhoI; George Chuck dissected immature tassels between 1mm and 3mm. Sharon Stanfield prepared the cDNA library in HybrizAP. Sample insert size range was 350 bp to 3 Kb with a 1 Kb average."	
ORIGIN	
Query Match	20.8%; Score 576.8; DB 13; Length 610;
Best Local Similarity	98.3%; Pred. No. 2.8e-49;
Matches 594; Conservative 0; Mismatches 7; Indels 3; Gaps 1;	
QY	1075 GCGGCGGAGGACACTCCGCGCAGCGCGGCGCTGCCGCTCCGGCCCAAGGAAGGACC 1134
Db	1 GCGGCGGAGGACACTCGCGCAGCGCGCAGCGCTGCCGCTCCGGCCCAAGGAAGGACC 60
QY	1135 TGCAATGCTGCTTGAGGATCCAGCGCTGCCCGCTGTCGACGCGCGCGCGTGCACG 1194
Db	61 TGCAATGCTGCTTGAGGATCCAGCGCTGCCCGCTTCGAGCGCGCGCGCTGCACG 120
QY	1195 TCCTTCGCGCGCGCGCGCTGACCATCGCGACGTCCTCGCCARAGGAGCCACAGGCTTACG 1254
Db	121 TCCTTCGCGCGCGCGCGCTGACCATCGCGAGTCCTTCGCCAAGGAGCCACAGGCTTACG 180
QY	1255 ACGAGTACGGCGCGCAGCACTACAGCAGCAGGACGAAGAACGGGAGCGCGCGCGGACGA 1314
Db	181 ACGAGTACGGCGCGCAGCACTACAGCAGCAGGACGAAGAACGGGAGCGCGCGCGGACGA 240
QY	1315 AGGCGGCGCGCAGCCTGTCGNAGCTGGGTCCACTCGAGCGCGCAGCTGTACCCCAAGG 1374
Db	241 AGGCGGCGCGCAGCCTGTTCGAAGCTGGGTCCACTCGAGCGCGCAGCTGTACCCCAAGG 300
QY	1375 ACGACGCGCAGGGAGGCGCGCGGTGGCGATTCGCGCGCGCGCAGCGGTGATGACGCGGC 1434
Db	301 ACGACGCGGAGGGAGGACGCGC---GGCGATCGCGCGCGCGCGGTGATGACGCGGC 357
QY	1435 TCATCCTCATCATGGTGTGAGGAAGCTGATCCGGAAACCCCAACACTACTCCAGGCTCA 1494
Db	358 TCATCCTCATCATGGTGTGAGGAAGCTGATCCGGAAACCCCAACACTACTCCAGGCTCA 417
QY	1495 TCGGCGTCTGCTGCTCCCTGGTCTCTACAGGTGGGCATCGAGATGCGCAGCGATCATCG 1554
Db	418 TCGGCGTCTGCTGCTCCCTGGTATCTTACAGGTGGGCATCGAGATGCCAGCGATCATCG 477
QY	1555 CCCGTCGATTTTCGATCTCTGTCGAGCAGCGGCTCTCGGGAATGGCCATGTTTACGCTAGGCC 1614



244	Db	CGACGGCGAGGGGAGGCAAGCGCGC---GGCGATGCGCGCGGCGAGCGTGATGACGGGCT	300
1436	Qy	CATCCTCATCATGCTGTGGAGGAAGCTGATCCGGAAACCCAAACACCTACTTCCAGCCTCAT	1495
301	Db	CATCCTCATCATGCTGTGGAGGAAGCTGATCCGGAAACCCAAACACTTACTCCAGCCTCAT	360
1496	Qy	CGCGTTCGTTCGTTCCTGTCTCCTACAGTGGGGGATCGAATGCGACGAGATCATCGC	1555
361	Db	CGCGTTCGTTCGTTCCTGTCTCCTACAGTGGGGGATCGAATGCGACGAGATCATCGC	420
1556	Qy	CCGGTCGATTTTCGATCTCTGTCGGACGGGGTCTCGGATCGCCATGTTCAAGCCTAGGCGCT	1615
421	Db	CCGGTCGATTTTCGATCTCTGTCGGACGGGGTCTCGGATCGCCATGTTCAAGCCTAGGCGCT	480
1616	Qy	GTTTCATGGCGGTGACGCCGAGGATCATTCGGTTCGGGGAAACAAGTGGCGGCCATTCGGCAT	1675
481	Db	GTTTCATGGCGGTGACGCCGAGGATCATTCGGTTCGGGGAAACAAGTGGCGGCCATTCGGCAT	540
1676	Qy	GGGGGTCGGGTTCGTTCGACAGCCCCCGGGTTCATGGCCGCGCGCTCCATCGCCCGTCGGTCT	1735
541	Db	GGGGGTCGGGTTCGTTCGACAGCCCCCGGGTTCATGGCCGCGCGCTCCATCGCCCGTCGGTCT	600
1736	Qy	GGCGGGCGTCTCTCTCCACATCGCCATCGT	1765
601	Db	GGCGGGCGTCTCTCTCCACATCGCCATCGT	630

RESULT 12  
BU037419  
LOCUS  
DEFINITION  
BU037419 627 bp mRNA linear EST 23-AUG-2002  
946138F12.Y1 946 - tassal primordium prepared by schmidt lab Zea  
mays cDNA. mRNA sequence.

ACCESSION	BU037419	
VERSION	BU037419.1	GI:22472939
KEYWORDS	EST.	
SOURCE	Zea mays	
ORGANISM	Zea mays	
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea. (bases 1 to 627)	

AUTHORS	Walbot, V.
TITLE	Maize ESTs from various cDNA libraries sequenced at Stanford University
JOURNAL	Unpublished (1999)
COMMENT	Contact: Walbot V Department of Biological Sciences Stanford University 855 California Ave, Palo Alto, CA 94304, USA Tel: 650 723 2227 Fax: 650 725 8221 Email: walbot@stanford.edu Plate: 946138 row: F column: 12.

```

FEATURES
    source
        Location/Qualifiers
            1. 627
                /organism="Zea mays"
                /mol_type="mRNA"
                /cultivar="OH43"
                /db_xref="taxon:4577"
                /tissue_type="tassels"
                /dev_stage="just after the transition from vegetative to
                inflorescence development"
                /lab_host="XLOLR"
                /clone_lib="946 - tassel primordium prepared by Schmidt
                lab"
                /note="Organ: tassels; Vector: HybrizAP; Site 1: EcoRI;
                Site 2: XhoI; George Chuck dissected immature tassels
                between 1mm and 3mm. Sharon Stanfield prepared the cDNA
                library in HybrizAP. Sample insert size range was 350 bp
                to 1 kb with a 1 kb average."

```

## ORIGIN

	Query Match	20.5%;	Score 569,	DB 13;	Length 627;
	Best Local Similarity	97.0%;	Pred. No. 1.7e-48;		
	Matches 604;	Conservative 0;	Mismatches 10;	Indels 9;	Gaps 2;
Qy	989	CGCGGACTTCTTCAACATCGTTCGGCGCCGCGCCGCAAGGAGGCGGAGGAGCGCGCGGGGA	1048		
Db	14	CGCGGACTTCTTCAACATCGTCTTTTCGCGCCGCAAGGAGGCGGAGGAGCGCGCGGGGA	73		
Qy	1049	CGAGGAGAAAGGGCGCATCGGCGGCGCGCGCGGAGGACACTCGCCGAGCGCGCAGAGCCGT	1108		
Db	74	CGAGGAGAAAGGGCGCAT-----GCGGCGCGGAGGACACTCGCCGAGCGCGCAGAGCCGT	127		
Qy	1109	CGCGGTCGCGCCCAAGAGAAAGCACTTGACATGCTCTGAGTCTCCAGCGCCTCGCC	1168		
Db	128	CGCGGTCGCGCCCAAGAGAAAGCACTTGACATGCTCTGAGTCTCCAGCGCCTCGCC	187		
Qy	1169	CGTGTCCGAGCGCGCGCGCTGCACTTTCGCGCGCGCGCGCGCTGACCATGCCGACGT	1228		
Db	188	CGTCTCCGAGCGCGCGCGCTGACCTCTTTCGCGCGCGCGCGCTGACCATGCCGACGT	247		
Qy	1229	CCTCGCCAAAGGAGCCAGGCTTACGACGATACGCGGCGGACGACTACAGCAGCAGGAC	1288		
Db	248	CCTCGCCAAAGGAGCCAGGCTTACGACGATACGCGGCGGACGACTACAGCAGCAGGAC	307		
Qy	1289	GAAGAACGGGAGCGCGCGCGGACAAAGGCGCGGACGCTGTCGAAAGTGTGGGTCCAA	1348		
Db	308	GAAGAACGGGAGCGCGCGCGGACAAAGGCGCGGCGGCGGACGCTGTCGAAAGTGTGGGTCCAA	367		
Qy	1349	CTCGAGCGCGAGCTGTACCCCAAGGACGACGCGGAGGAGGCGCGCGGTGGCGAT	1408		
Db	368	CTCGAGCGCGAGCTGTACCCCAAGGACGACGCGGAGGAGGCGCGCGGTGGCGAT	424		
Qy	1409	GCGCCGCGGAGCGTGATGACGCGGCTCATCTCATCTGTTGTGAGGAAGTGTATCCG	1468		
Db	425	GCGCCGCGGAGCGTGATGACGCGGCTCATCTCATCTGTTGTGAGGAAGTGTATCCG	484		
Qy	1469	GAACCCCAACACTACTCCAGCCTCATCGCGCTCGTCTGGTCCCTGGTCTCTACAGGTG	1528		
Db	485	GAACCCCAACACTACTCCAGCCTCATCGCGCTCGTCTGGTCCCTGGTCTCTACAGGTG	544		
Qy	1529	GGGCATCGAGTCCAGCGCATCATCGCCGGTTCGATTTTCGATCCTGTCCGACCGCGGTCT	1588		
Db	545	GGGCATCGAGTCCAGCGCATCATCGCCGGTTCGATTTTCGATCCTGTCCGACCGCGGTCT	604		
Qy	1589	CGGGATGGCCATGTTTCAGGCTAG	1611		
Db	605	CGGGATGGCCATGTTTCAGGCTAG	627		

RESULT 13					
CG456385					
LOCUS	794 bp	DNA	linear	GSS 17-SEP-2003	
DEFINITION	PUIKJ57TBB ZM 0.6_1.0 KB Zea mays genomic clone ZMMeta0603118, genomic survey sequence.				

CG456385	
CG456385.1	GI:34841385
GSS.	
Zea mays	
Zea mays	
ORGANISM	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
REFERENCE	
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Zea. (bases 1 to 794)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

1 (bases sent to 794)  
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,  
Resnick, A., Fraser, C.N., Yuan, Y., San Miguel, P., Ma, J. and  
Bennetzen, J.  
Maize Genomics Consortium  
Unpublished (2003)  
Other GSS: FURKJ57DB  
Contact: Cathy Whitelaw  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843

Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: TR  
Class: sheared ends.

FEATURES  
source Location/Qualifiers  
1..794  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone="ZMMB7a0603118"  
/clone\_lib="ZM 0.6-1.0 kb"  
/note="Vector: pCR4-TOP0; Site 1: EcoRI; 0.6-1.0 kb high  
Cor selected genomic DNA library"

ORIGIN  
Query Match 20.5%; Score 568.4; DB 29; Length 794;  
Best Local Similarity 88.8%; Pred. No. 1.8e-48;  
Matches 673; Conservative 0; Mismatches 1; Indels 84; Gaps 2;  
QY 1608 CTAGGCTGTTCATGGCGCTCGACGAGGATCATCGGTGCGGGAACAAGCTGGCGGCC 1667  
Db 37 CCAGGCTGTTCATGGCGCTCGACGAGGATCATCGGTGCGGGAACAAGCTGGCGGCC 96  
QY 1668 ATCCGATGGCGCTCGGTTCGTCGAGGCCCGCGGTTCATGGCGCGCTCCATCGCC 1727  
Db 97 ATCCGATGGCGCTCGGTTCGTCGAGGCCCGCGGTTCATGGCGCGCTCCATCGCC 156  
QY 1728 GTCGGTCTGCGCGCTCTCTCCACATCGCCATCG----- 1764  
Db 157 GTCGGTCTGCGCGCTCTCTCCACATCGCCATCGCCATCGCGCGCTGTGGCTT 216  
QY 1765 ----- 1764  
Db 217 GACGTAAACGGTGGATGTGATGGAAACAGGTAGGACCAACATTCGTTTCCCTTTCTT 276  
QY 1765 TCCAGGCTGTCTGCTCCTCAGGGGATCGCGGTTCGTTTCGCAAGAGTAGCGGCTTC 1824  
Db 277 TCCAGGCTGTCTGCTCCTCAGGGGATCGTCGCTTCGTTTCGCAAGAGTAGCGGCTTC 336  
QY 1825 ATCCGATCTGAGCACAGGATATGGTTCGAATAAATGTCATGCTATGTTTCACTATGAT 1884  
Db 337 ATCCGATCTGAGCACAGGATATGGTTCGAATAAATGTCATGCTATGTTTCACTATGAT 396  
QY 1885 ATAGTTAACGGGAAACAGAAAGCAATTCGATGACGACACTGAAATTCATATGAT 1944  
Db 397 ATAGTTAACGGGAAACAGAAAGCAATTCGATGACGACACTGAAATTCATATGAT 456  
QY 1945 TCATTAATAATGATGGTGTGTTTCATGACGTGACGTCAAGAACCACTAATAAGCACTGAT 2004  
Db 457 TCATTAATAATGATGGTGTGTTTCATGACGTGACGTCAAGAACCACTAATAAGCACTGAT 516  
QY 2005 CTAGGACAGCATCAGCATGATGATGCTGTTTCTCCTGACAACTCTGATTTCTTACT 2064  
Db 517 CTAGGACAGCATCAGCATGATGATGCTGTTTCTCCTGACAACTCTGATTTCTTACT 576  
QY 2065 ACAGAGTGTGCTTCACTCATCCATCCAGATGATCATAACACTACTGATGATC-TTT 2123  
Db 577 ACAGAGTGTGCTTCACTCATCCATCCAGATGATCATAACACTACTGATGATC-TTTT 636  
QY 2124 TTTTGTGATTTCTGTCAGCGTATCTTCGGGATGCTATGCTCTGCCCATCACCCTGG 2183  
Db 637 TTTTGTGATTTCTGTCAGCGTATCTTCGGGATGCTATGCTCTGCCCATCACCCTGG 696  
QY 2184 TCTACTACATCTGTCGGGCTGTGAGCCTCTCTCGCTCGCTTCTTCAGCGGTGCGGAAGG 2243  
Db 697 TCTACTACATCTGTCGGGCTGTGAGCCTCTCTCGCTCGCTTCTTCAGCGGTGCGGAAGG 756  
QY 2244 CGCCATGCTGTGGTATTCGGCCCAATGAATTTCT 2281  
Db 757 CGCCATGCTGTGGTATTCGGCCCAATGAATTTCT 794

RESULT 14  
CG317974

LOCUS  
DEFINITION

CG317974

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1..776

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

/db\_xref="taxon:4577"

/clone="ZMMB7a0570013"

/clone\_lib="ZM 0.7-1.5 kb"

/note="Vector: pBCSK; Site 1: HindII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

ORIGIN

Query Match 20.3%; Score 561; DB 29; Length 776;  
Best Local Similarity 87.4%; Pred. No. 1e-47;  
Matches 679; Conservative 0; Mismatches 98; Gaps 2;  
QY 795 GTCTCTCTCCCGCGGGGACCGCGGGGGTGCCTGTCGCGTCCGCAAGTCCACGAC 854  
Db 1 GTCTCTCTCCCGCGGGGACCGCGGGGGTGCCTGTCGCGTCCGCAAGTCCACGAC 60  
QY 855 TCGCGTCCGAGCGCGCTGCTCGCACTCGCACTCCAGACCATCGAGCCCGGTGTCTCC 914  
Db 61 TCGCGTCCGAGCGCGCTGCTCGCACTCGCACTCCAGACCATCGAGCCCGGTGTCTCC 120  
QY 915 AACCTCTCCCGGTGGAGATCTACTCGGTGCACTGCTCGCAACCCACCCCGCGCGG 974  
Db 121 AACCTCTCCCGGTGGAGATCTACTCGGTGCACTGCTCGCAACCCACCCCGCGCGG 180  
QY 975 TCAGGTTCAACCAACCGGACCTTCTTCAACATCTGTCGGCGCGCGCCGCAAGGAGCGGA 1034  
Db 181 TCAC-CTTCAACCAACCGGACCTTCTTCAACATCTGTCGGCGCGCGCCGCAAGGAGCGGA 239  
QY 1035 GGAGCGCGGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1094  
Db 240 GGAGCGCGGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 299  
QY 1095 CAGCGCGAGCGCGCTCGCGTCCGCGCCCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1154  
Db 300 CAGCGCGAGCGCGCTCGCGTCCGCGCCCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 359  
QY 1155 TCCAGCGCTCCCGCTGTCGAGCGCGCGCGCGCTTTCGGCGCGCGCGCGCT 1214  
Db 360 TCCAGCGCTCCCGCTGTCGAGCGCGCGCGCGCTTTCGGCGCGCGCGCGCT 419  
QY 1215 GACCATGCGCGCTGCTCTCGCCAA----- 1237

CG317974 776 bp DNA linear GSS 26-AUG-2003  
OGWGA91TV ZM\_0.7-1.5 KB Zea mays genomic clone ZMMB7a0570013,  
genomic survey sequence.  
CG317974  
CG317974.1 GI:34235240  
GSS.  
Zea mays  
Zea mays  
Zea mays

Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 776)  
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,  
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,  
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.  
Consortium for Maize Genomics  
Unpublished (2002)  
Other\_GSSs: OGWGA91TV  
Contact: Cathy Whitelaw  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: TR  
Class: sheared ends.

Location/Qualifiers  
1..776  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone="ZMMB7a0570013"  
/clone\_lib="ZM 0.7-1.5 kb"  
/note="Vector: pBCSK; Site 1: HindII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

[illegible]





Db 335368 GCGGCGCGGAGGAGCATGCGGCAATGCGCGCGCGCGCGTCCCGCGCGGCGGAC 335427  
QY 943 TGAGTGTGCGGCAACCCACCGCGCGGCTCAGCTTCAACACGCGGACTTCTTCA 1002  
Db 335428 AGGCTGTGTCGCGCGCTCGCGCTTCCCGCTGCGCGGAGACAGCGCGCGCGCGCG 335487  
QY 1003 ACATCGTCG 1062  
Db 335488 GCGCGCGCGGACTCG 335547  
QY 1063 CATGCG 1122  
Db 335548 GCGCGCGGACAGCG 335607  
QY 1123 AGAGGAAGGACCTGACATGCTCGTCTGAGGTCCTCAGCGCGCTCGCGCGCGCG 1182  
Db 335608 CCGGATCG 335667  
QY 1183 CCGCGCTGCG 1242  
Db 335668 CCGCGGAGTGCACTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 335727  
QY 1243 CCGAGCGCTAGCAGGAGTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1302  
Db 335728 AGCAGCGCGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 335787  
QY 1303 GCGCGCGGCAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1358  
Db 335788 AGCG 335847  
QY 1359 CAGCTGTATCCCAAGGACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1418  
Db 335848 CCGCGCGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 335907  
QY 1419 AGCGTGATGACGCGG 1433  
Db 335908 CCGATGAGCGCGCGCG 335922

RESULT 4  
US-09-103-840A-1  
; Sequence 1, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; FILE REFERENCE: 24366-20007.00  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; OTHER INFORMATION: H37Rv  
US-09-103-840A-1

Query Match 3.6%; Score 100.2; DB 3; Length 4411529;  
Best Local Similarity 46.3%; Pred. No. 5.4e-09;  
Matches 368; Conservative 0; Mismatches 423; Indels 4; Gaps 1;  
QY 643 ACCG 702  
Db 335020 ATCCCGCGCGGGAAGCCCTGCGCAACTCGCGCGCGCGCGCGCGCGCGCG 335079  
QY 703 TCTCCTCTCGCGCTCGACTCCGACGCTCGCTCGCGCGCGCGCGCGCGCGCG 762

Db 335080 ACCGTCCCGCGCGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 335139  
QY 763 CCAGAGCCCGACGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 822  
Db 335140 GCGCGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 335199  
QY 823 GGGTGGCGCTCAGCGTGCAGAGTCCAGAGTCCAGAGTCCAGAGTCCAGAGTCCAG 882  
Db 335200 CCGCTATATCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 335259  
QY 883 CGCACTCCAGAGCAATGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 942  
Db 335260 GCGCGCGCGGAGCGGAGCATGCGGCAATGCGCGCGCGCGCGCGCGCGCGCG 335319  
QY 943 TGAGTGTGTCG 1002  
Db 335320 AGGCTGTGTCG 335379  
QY 1003 ACATCGTCG 1062  
Db 335380 GCGCGCGCGACTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 335439  
QY 1063 CATGCG 1122  
Db 335440 GCGCGCGAAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 335499  
QY 1123 AGAGGAAGGACCTGACATGCTCGTCTGAGGTCCTCAGCGCGCTCGCGCGCG 1182  
Db 335500 CCGGATCG 335559  
QY 1183 CCGCGCTGCACTGCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1242  
Db 335560 CCGCGGAGTGCACTGCGCTGCTGTAATCGCGCGCGCGCGCGCGCGCGCG 335619  
QY 1243 CCGAGCGCTAGCAGAGTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1302  
Db 335620 AGCAGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 335679  
QY 1303 GCGCGCGCGGACGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1358  
Db 335680 AGCG 335739  
QY 1359 CAGCTGTATCCCAAGGACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1418  
Db 335740 CCGCGCGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 335799  
QY 1419 AGCGTGATGACGCGG 1433  
Db 335800 CCGATGAGCGCGCGCG 335814

RESULT 5  
US-09-103-840A-1/c  
; Sequence 1, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; FILE REFERENCE: 24366-20007.00  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 4411529  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; OTHER INFORMATION: H37Rv  
US-09-103-840A-1

```
Query Match 3.6%; Score 99.2; DB 3; Length 4411529;
Best Local Similarity 46.4%; Pred. No. 8.3e-09;
Matches 359; Conservative 0; Mismatches 413; Indels 2; Gaps 1;

QY 642 TACCGCGCGCGCGCGCGTCTCTCTCGACCAAGTTCCTCCGACGGCGCGCGCGTCCATC 701
DB 3930731 TACGGTGATCGCGGTCAACCTACGCGACGCGCGCGCGCGCGCGCGCGCGCGCG 3930672
QY 702 GTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 761
DB 3930671 GCGCGTGCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3930612
QY 762 GCGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 821
DB 3930611 GCGCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3930552
QY 822 CGGCTGCGCGTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 881
DB 3930551 CCGCGCGCGTGCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3930492
QY 882 TCGCACTCTCCAGACCATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 941
DB 3930491 GCGCGCGCGGTAAACGCGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3930432
QY 942 CTGCGAGTCTGCGCGCAACCCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1001
DB 3930431 GCGCGTGCCTGCGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3930372
QY 1002 AACATCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1061
DB 3930371 GCGCGCGTTCGCGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3930312
QY 1062 GATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1121
DB 3930311 GTTGCCTGCGCGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3930252
QY 1122 AAGAGGAAGACTGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1181
DB 3930251 GCGGAACCGCGCGTGTATCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3930192
QY 1182 GCGCGCGTGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1241
DB 3930191 GCGCGCGTGCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3930132
QY 1242 GCGCGCGTGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1301
DB 3930131 GCGCGTGCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3930072
QY 1302 GCGCGCGCGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1361
DB 3930071 GGTGCGCGTGGCACCGCGCG--CGATGTTGCGCGCGTTCGCGCGCGCGCG 3930014
QY 1362 CTGTACCCAGACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1415
DB 3930013 CCGTTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3929960
```

```
RESULT 6
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: PRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
```

```
; SOFTWARE: PatentIn Ver. 2.1.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2
```

```
Query Match 3.5%; Score 98; DB 3; Length 4403765;
Best Local Similarity 47.1%; Pred. No. 1.4e-08;
Matches 369; Conservative 0; Mismatches 410; Indels 5; Gaps 2;

QY 635 CTTCGAGTACCGCGCGCGCGCGCGTCTCTCTCGACCAAGTTCCTCCGACGGCGCGCGCG 694
DB 3924520 CTACGGTGATCGCGGTCAACCTACGCGACGCGCGCGCGCGCGCGCGCGCGCGCG 3924461
QY 695 GTCCATCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 754
DB 3924460 CGCGCGTGCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3924401
QY 755 GCTCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 814
DB 3924400 CGCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3924341
QY 815 CGCGCGCGCGGTGC---GGGTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 871
DB 3924340 ACCCGCGCGTGCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3924281
QY 872 GTGCTGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 931
DB 3924280 CGCGCGCGCGGTAAACGCGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3924221
QY 932 GATCTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 991
DB 3924220 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3924161
QY 992 CGATCTCTTCAACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1051
DB 3924160 AGCGCGCGGTGCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3924101
QY 1052 GGAGAAAGGCGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1111
DB 3924100 CGGTGCGCGCGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3924041
QY 1112 CGTGCAGGCGCGCAAGAGAGAAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1171
DB 3924040 CGCGCGTCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3923981
QY 1172 GTCCGAGCGCGCGCGCGCGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1231
DB 3923980 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3923921
QY 1232 CGCCAAAGAGCGCGCGCGCGCGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1291
DB 3923920 CGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3923861
QY 1292 GAACGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1351
DB 3923860 CGTTGCGCGCGGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3923803
QY 1352 GACGCGCGAGTCTTACCCCAAGGAGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1411
DB 3923802 GCGCGCGAGCGCGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3923743
QY 1412 GCGC 1415
DB 3923742 GCGC 3923739
```

RESULT 7



US-09-249-585A-2/c  
; Sequence 2, Application US/09249585A  
; Patent No. 6417002

; GENERAL INFORMATION:  
; APPLICANT: Horlick, Robert

; TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISODES  
; FILE REFERENCE: 0867/0905

; CURRENT APPLICATION NUMBER: US/09/249,585A  
; CURRENT FILING DATE: 1999-02-11

; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 2

; LENGTH: 1926

; TYPE: DNA

; ORGANISM: Epstein Barr Virus

; FEATURE:  
; NAME/KEY: CDS

; LOCATION: (1)..(1926)

; OTHER INFORMATION: coding strand of EBNA-1 DNA  
US-09-249-585A-2

Query Match 3.4%; Score 92.8; DB 4; Length 1926;  
Best Local Similarity 47.1%; Pred. No. 1.6e-08;

Matches 362; Conservative 0; Mismatches 392; Indels 14; Gaps 2;

QY 122 CGCGACCCCTCTCCCTCCCGGTCGCCGAGCAATCGCGGCATGATCACGCG 181  
DB 1061 CGCGGCGCTCCACTACTCTCGACCCGGGCTCACTACTCTCGACCCGGGCTCCA 1002

QY 182 GCTGACCTCTACCAAGTGTGACGGGGTGGTGCCTGTAAGTGCCATGAGCTGGC 241  
DB 1001 CTGCTCTCTGACCCGGGCTCCACTCTCTGCCCCCTCTGCTGCTGCCCCCTCT 942

QY 242 GTACGGCTCCGTGCGCTGTCGCGATCTTACCGCGGACAGTGTCCGGGATCAACG 301  
DB 941 CCGTCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCGG 882

QY 302 CTTCGTGCGCTCTTCGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 361  
DB 881 CTTCTGCGGCT 822

QY 362 CTTGCGGATGAACTGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 421  
DB 821 CTTGCGGATGAACTGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 762

QY 422 GCTGTCGCGTGGGCTCTCCGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 481  
DB 761 CTTGTCGCGTGGGCTCTCCGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 702

QY 482 CATCAGCTCTTCTCCCTC-----TCCAGCTCTCCCAACAGCTCTGTCATGGG 529  
DB 701 CTTCTGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCG 642

QY 530 CATCCGGTCTGCGAGGATGAGGCGCTGCTGCGCGGCGAGCTCATGGTCCAGGT 589  
DB 641 CTTGCGGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCG 582

QY 590 CTTGCTCTCTCCAGTGCATCATCTGTATACGCTCATGCTCTCTCTCTCTCTCTCT 649  
DB 581 CTTGCTCTCTCCAGTGCATCATCTGTATACGCTCATGCTCTCTCTCTCTCTCTCT 522

QY 650 CCGCGCGGCTCTGCTCTGAGGATGAGGCTCTGCGGCGGCGGCTCTGCGGCGGCG 709  
DB 521 CTTGCT 464

QY 710 CCGCGTCTGACTCCGAGTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCT 769  
DB 463 CTTGCGGCT 404

QY 770 CGACGGGCTCTGCGGCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 829  
DB 403 CTTGCT 344

QY 830 CTTACCGTCTGCGCAAGTCCACCAAGTCTGCGCTCCGAGCGCGCTGCTC 877  
DB 343 CCGCT 296

## RESULT 8

US-09-410-399-3/c

; Sequence 3, Application US/09410399

; Patent No. 6482587

; GENERAL INFORMATION:

; APPLICANT: Robertson, Erle S.

; APPLICANT: Cottler, Murray A.

; TITLE OF INVENTION: Methods to Inhibit or Enhance the Binding of Viral DNA

; TITLE OF INVENTION: to Genomic Host DNA

; FILE REFERENCE: UM-03778

; CURRENT APPLICATION NUMBER: US/09/410,399

; CURRENT FILING DATE: 1999-10-01

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 3

; LENGTH: 1926

; TYPE: DNA

; ORGANISM: Epstein-Barr virus

US-09-410-399-3

Query Match 3.4%; Score 92.8; DB 4; Length 1926;

Best Local Similarity 47.1%; Pred. No. 1.6e-08;

Matches 362; Conservative 0; Mismatches 392; Indels 14; Gaps 2;

QY 122 CGCGACCCCT 181  
DB 1061 CGCGGCGCTCCACTACTCTCTGACCCGGGCTCCACTACTCTCTGACCCGGGCTCCA 1002

QY 182 GCTGACCTCTACCAAGTGTGACGGGGTGGTGCCTGTAAGTGCCATGAGCTGGC 241  
DB 1001 CTGCTCTCTGACCCGGGCTCCACTCTCTGCCCCCTCTGCTGCTGCCCCCTCT 942

QY 242 GTACGGCTCCGTGCGCTGTCGCGATCTTACCGCGGACAGTGTCCGGGATCAACG 301  
DB 941 CCGTCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCGG 882

QY 302 CTTCGTGCGCTCTTCGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 361  
DB 881 CTTCTGCGGCT 822

QY 362 CTTGCGGATGAACTGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 421  
DB 821 CTTGCGGATGAACTGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 762

QY 422 GCTGTCGCGTGGGCTCTCCGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 481  
DB 761 CTTGTCGCGTGGGCTCTCCGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 702

QY 482 CATCAGCTCTTCTCCCTC-----TCCAGCTCTCCCAACAGCTCTGTCATGGG 529  
DB 701 CTTCTGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCG 642

QY 530 CATCCGGTCTGCGAGGATGAGGCGCTGCTGCGCGGCGAGCTCATGGTCCAGGT 589  
DB 641 CTTGCGGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCG 582

QY 590 CTTGCTCTCTCCAGTGCATCATCTGTATACGCTCATGCTCTCTCTCTCTCTCTCT 649  
DB 581 CTTGCTCTCTCCAGTGCATCATCTGTATACGCTCATGCTCTCTCTCTCTCTCTCT 522

QY 650 CCGCGCGGCTCTGCTCTGAGGATGAGGCTCTGCGGCGGCGGCTCTGCGGCGGCG 709  
DB 521 CTTGCT 464

QY 710 CCGCGTCTGACTCCGAGTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCT 769  
DB 463 CTTGCGGCT 404



```

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 2580 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: unknown
;   TOPOLOGY: unknown
; MOLECULE TYPE: DNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-359-081-2

Query Match          3.4%; Score 92.8; DB 4; Length 2580;
Best Local Similarity 47.1%; Pred. No. 1.7e-08;
Matches 362; Conservative 0; Mismatches 392; Indels 14; Gaps 2;

Qy 122 CGCGACCCCTCTCCCTCCCGGCTCCCGGACCAAGCCATCCCGGCGCATGATCAACGC 181
Db 1444 CGGCGGCTCCACTACCTCCCTCGACCGCGGCTCCACTACCTCTCGACCGCGGCTCCA 1385

Qy 182 GCTGGACCTTACCAAGTGTGACGGGGTGGTGGCTGTACGTGGCCATGAGCTGGC 241
Db 1384 CTGCTCTCTGACCGCGGCTCCACCTCTCTGCTCTGCTGCTGCTGCTGCTGCTGCT 1325

Qy 242 GTACGGCTCCGCTCGGCTGTGGCGCATCTTACACCGCGGACAGTGTCTCCGGGATCAACCG 301
Db 1324 CTGCTCTGCGGCTCTGCGGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 1265

Qy 302 CTTCGTGGCGCTCTTCGGCGGTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 361
Db 1264 CTTCTGCGGCTCTGCTCTGCGGCTCTGCGGCTCTGCTCTGCTCTGCTCTGCTCTGCT 1205

Qy 362 CTTCGCCATGAACCTGCGCTTCTTCGGCGCGGACAGCTGTGAGAAAGTGGCGGCTCTGCG 421
Db 1204 CTCTCTCTGCTCTGCGGCTCTGCGGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 1145

Qy 422 GCTGTGGCGTGTGGCTCTCCCGGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 481
Db 1144 CTGCGGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCTCTGCTCTGCTCTGCTCTGCT 1085

Qy 482 CATCAAGCTCTTCTCCCTC-----TCCAGCTCTCCCAACAGCTCTGCTCATGGG 529
Db 1084 CTCTCTCTGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCTCTGCTCTGCTCTGCTCTGCT 1025

Qy 530 CATCCCGTGTGGAGGATGTACGGCGGTGTGCGCGGCGGACGCTCATGCTGCTCAAGT 589
Db 1024 CTGCGGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCTCTGCTCTGCTCTGCTCTGCT 965

Qy 590 GTGTCTCTCAGTGATCATGTGTACAGCTCATGCTCTTCTCTCTCTCTCTCTCTCTCTCT 649
Db 964 CTGCGGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCTCTGCTCTGCTCTGCTCTGCT 905

Qy 650 CGCGGCGGCTCTGCTCTGAGTTCGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 709
Db 904 CTCTCTCTGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCTCTGCTCTGCTCTGCTCTGCT 847

Qy 710 CGCGTGTAGTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 769
Db 846 CTGCGGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 787

Qy 770 CGAGCGGCTCTGCGGCGGCGGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 829
Db 786 CTCTCTCTGCTCTGCGGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 727

Qy 830 CGTCAACGTTGGCAAGTCCACAGTCTGCGTCTCGAGCGCGCGCTGCTC 877
Db 726 CCCCTCTCTGCTCTGCGGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 679

RESULT 11
US-09-130-114-1
; Sequence 1, Application US/09130114
; Patent No. 5976807
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert A.
```

```

; APPLICANT: Damaj, Bassem B.
; APPLICANT: Robbins, Alan K.
; TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes
; TITLE OF INVENTION: From Multiple Transfected Episomes
; FILE REFERENCE: 0867/1D9030S1
; CURRENT APPLICATION NUMBER: US/09/130,114
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 5452
; TYPE: DNA
; ORGANISM: VEBNA
US-09-130-114-1

Query Match          3.4%; Score 92.8; DB 2; Length 5452;
Best Local Similarity 47.1%; Pred. No. 2.1e-08;
Matches 362; Conservative 0; Mismatches 392; Indels 14; Gaps 2;

Qy 122 CGCGACCCCTCTCTCTCTCCCGGCTCCCGGACCAAGCCATCCGGGCGCATGATCAACGC 181
Db 1361 CGGCGGCTTCCACTACCTCTCTGACCGCGGCTCCACTACCTCTCGACCGCGGCTCCA 1420

Qy 182 GCTGGACCTTACCAAGTGTGACGGCGGTGTGCGGCTGTACGTGGCCATGACGCTGCG 241
Db 1421 CTGCTCTCTGACCGCGGCTCCACTCTCTGTCTCTGCTCTCTCTCTCTCTCTCTCTCT 1480

Qy 242 GTACGGCTCTGCTCGGCTGTGGCGCATCTTACAGCGCGGACAGTGTCTCGGGATCAACCG 301
Db 1481 CTGCTCTCTGCGGCTCTGCGGCTCTGCTCTCTGCTCTCTGCGGCTCTCTCTCTCTCT 1540

Qy 302 CTTCGTGGCGCTTTCGGCGGTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 361
Db 1541 CTCTCTGCGGCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1600

Qy 362 CTTCGCCATGAACCTGCGCTTCTTCGGCGCGGACAGCTGTGAGAAAGTGGCGGCTCTGCG 421
Db 1601 CTCTCTCTGCTCTGCGGCTCTGCGGCTCTGCTCTCTGCTCTCTGCGGCTCTCTCTCT 1660

Qy 422 GCTGTGGCGGCTGCGGCTCTCCCGGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 481
Db 1661 CTGCGGCTCTGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1720

Qy 482 CATCAAGCTCTTCTCCCTC-----TCCAGCTCTCCCAACAGCTCTGCTCATGGG 529
Db 1721 CTCTCTGCTCTGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1780

Qy 530 CATCCCGTGTGGAGGATGTACGGCGGTGTGCGCGGCGGACGCTCATGCTGCTCAAGT 589
Db 1781 CTGCGGCTCTGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1840

Qy 590 CGTGTCTCTTCCAGTGTATCATGTGTACAGCTCATGCTCTTCTCTCTCTCTCTCTCTCT 649
Db 1841 CTGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1900

Qy 650 CGCGGCGGCTCTGCTCTGAGTTCGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCTCT 709
Db 1901 CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1958

Qy 710 CGCGGCTCTGAGTTCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 769
Db 1959 CTGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2018

Qy 770 CGAGCGGCTCTGCGGCGGCGGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 829
Db 2019 CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2078

Qy 830 GGTCAACGTTGGCAAGTCCACAGTCTGCGGCTCCGAGGCGCGGCTGCTC 877
Db 2079 CCCCTCTCTGCTCTGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2126

RESULT 12
```

US-09-647-344A-14  
; Sequence 14, Application US/09647344A  
; Patent No. 6586180  
; GENERAL INFORMATION:  
; APPLICANT: Ruffner, Duane E.  
; APPLICANT: Chen, Zhidong  
; TITLE OF INVENTION: Directed Antisense Libraries  
; FILE REFERENCE: T6678.PCT.US  
; CURRENT APPLICATION NUMBER: US/09/647,344A  
; CURRENT FILING DATE: 2000-12-04  
; PRIOR APPLICATION NUMBER: PCT/US99/06742  
; PRIOR FILING DATE: 1999-03-28  
; NUMBER OF SEQ ID NOS: 50  
; SEQ ID NO 14  
; LENGTH: 8705  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: pshuttle  
US-09-647-344A-14

Query Match 3.4%; Score 92.8; DB 4; Length 8705;  
Best Local Similarity 47.1%; Pred. No. 2.4e-08;  
Matches 362; Conservative 0; Mismatches 392; Indels 14; Gaps 2;  
  
QY 122 CGCGGACCCCTCTCTCCCTCCCGCGTCCCGGACGATCCGCGGCATGATCACGCG 181  
Db 7226 CGGGGGCTCCATCTCTCTGACCCCGGCTTCACTACCTCTCGACCCCGGCTCCA 7285  
  
QY 182 GCTGGACCTTACAGTGTGACGGGGTGTGGCGCTGTAGTGGCCATGACGCTGGC 241  
Db 7286 CTGCTCTCTGACCCGCGCTCCACCTCTCTGCTCTCTGCTCTCTCTCTCTCTCT 7345  
  
QY 242 GTACGGCTCGCTCGCTGTTGGGCGCATCTTACGCGGACGATGCTCCGGGATCAACG 301  
Db 7346 CTGCTCTCTGCTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7405  
  
QY 302 CTTCGTTGGCGCTCTTCGCGCTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 361  
Db 7406 CTTCTGCT 7465  
  
QY 362 CTTCGCTGATGACCTGCTCTCTGCGGCTCTGCGGCGGACGCTGCAAGGTGCGTCTCG 421  
Db 7466 CTTCT 7525  
  
QY 422 GCTGCTGCGCTGCGCTCTCCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 481  
Db 7526 CTTGCT 7585  
  
QY 482 CATCAGCTCTTCT 529  
Db 7586 CTTGCT 7645  
  
QY 530 CATCCGCTCTGCGAGGCTATGACGGCGGTGTGTTGGCGGACGCTCATGTTCCAGGT 589  
Db 7646 CTTGCT 7705  
  
QY 590 GGTGCTCTCTGAGTGTATCTGTTGACAGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 649  
Db 7706 CTTGCT 7765  
  
QY 650 CGCGCGGCGCT 709  
Db 7766 CTTCT 7823  
  
QY 710 CCGGCTGACATCCGACCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 769  
Db 7824 CTGCT 7883  
  
QY 770 CGAGGGGCTGCGCGGCGGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 829  
Db 7884 CTTCT 7943

QY 830 CTTACCGTGGCAAGTCCACAGCTCGGCTCCGAGCGCGCTGCTC 877  
Db 7944 CCCT 7991

## RESULT 13

US-08-910-647-1/c  
; Sequence 1, Application US/08910647  
; Patent No. 6251433  
; GENERAL INFORMATION:  
; APPLICANT: Zuckermann et al.  
; TITLE OF INVENTION: Compositions and Methods for  
; TITLE OF INVENTION: Polynucleotide Delivery  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street  
; CITY: Emeryville  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 94608-2916  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/910,647  
; FILING DATE:  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fujita, Sharon M.  
; REGISTRATION NUMBER: 38,459  
; REFERENCE/DOCKET NUMBER: 1218.002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (510) 923-2706  
; TELEFAX: (510) 655-3542  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9600 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-910-647-1

Query Match 3.4%; Score 92.8; DB 3; Length 9600;  
Best Local Similarity 47.1%; Pred. No. 2.5e-08;  
Matches 362; Conservative 0; Mismatches 392; Indels 14; Gaps 2;

QY 122 CGCGACCCCT 181  
Db 1490 CGGGGCTCTCACTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCA 1431  
  
QY 182 GCTGACCTCTTACACAGTGTGACGGGGTGTGGCTGTGCTGCTGCTGCTGCTGCTGCT 241  
Db 1430 CTGCTCTCTCTGACCCCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1371  
  
QY 242 GTACGGCT 301  
Db 1370 CTTGCT 1311  
  
QY 302 CTTGCTGCGCT 361  
Db 1310 CTTCT 1251  
  
QY 362 CTTGCTGCTGAACT 421  
Db 1250 CTTCT 1191  
  
QY 422 GCTGCTGCGCTGCGCT 481





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 3, 2004, 04:23:14 ; Search time 906 Seconds  
(without alignments)  
11025.470 Million cell updates/sec

Title: US-10-030-884-13

Perfect score: 2769

Sequence: 1 ccacgcctccgtgcgcct.....aaaaaaaaaaaaaaaaaaag 2769

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2353733 seqs, 180373377 residues

Total number of hits satisfying chosen parameters: 4707466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:\*

- 1: /cgm2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgm2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgm2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgm2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgm2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgm2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgm2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgm2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgm2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgm2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgm2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgm2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 13: /cgm2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgm2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgm2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgm2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 17: /cgm2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 18: /cgm2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	836.8	30.2	1926	15	US-10-260-238-565 Sequence 565, App
2	698	25.2	1759	15	US-10-260-238-30 Sequence 30, Appl
3	681.2	24.6	2276	12	US-10-425-114-1349 Sequence 1349, Ap
4	612	22.1	2415	12	US-10-425-114-16771 Sequence 16771, A
5	586.4	21.2	651	15	US-10-260-238-5337 Sequence 5337, Ap
6	451.2	16.3	2346	12	US-10-424-599-134028 Sequence 134028, A
7	450	16.3	1273	12	US-10-425-114-15739 Sequence 15739, A
8	427.8	15.4	1313	12	US-10-425-114-11838 Sequence 11838, A
9	426.2	15.4	1334	12	US-10-425-114-2319 Sequence 2319, Ap
10	426.2	15.4	1199	12	US-10-425-114-29155 Sequence 29155, A
11	425.6	15.4	1321	12	US-10-425-114-33502 Sequence 33502, A
12	412.8	14.9	1091	12	US-10-425-114-3558 Sequence 3558, Ap
13	370	13.4	702	15	US-10-260-238-5464 Sequence 5464, Ap
14	347.2	12.5	2222	9	US-09-887-576-628 Sequence 628, App
15	319.8	11.5	3020	12	US-10-424-599-45725 Sequence 45725, A

Sequence 848, App  
Sequence 848, App  
Sequence 4000, Ap  
Sequence 59026, A  
Sequence 847, App  
Sequence 847, App  
Sequence 104351, A  
Sequence 134029, A  
Sequence 86375, A  
Sequence 11117, A  
Sequence 134033, A  
Sequence 29226, A  
Sequence 10077, A  
Sequence 7737, Ap  
Sequence 45726, A  
Sequence 1305, Ap  
Sequence 1305, Ap  
Sequence 45727, A  
Sequence 59027, A  
Sequence 71900, A  
Sequence 62375, A  
Sequence 5138, Ap  
Sequence 33472, A  
Sequence 13889, A  
Sequence 18686, A  
Sequence 52693, A  
Sequence 30762, A  
Sequence 2884, Ap  
Sequence 15102, A  
Sequence 1, Appli

#### ALIGNMENTS

#### RESULT 1

US-10-260-238-565  
Sequence 565, Application US/10260238  
Publication No. US20040016025A1  
GENERAL INFORMATION:  
APPLICANT: Budworth, Paul R.  
APPLICANT: Moughamer, Todd G.  
APPLICANT: Briggs, Steven P.  
APPLICANT: Cooper, Bret  
APPLICANT: Glazebrook, Jane  
APPLICANT: Goff, Stephen A.  
APPLICANT: Katagiri, Fumiyaki  
APPLICANT: Kreps, Joel  
APPLICANT: Provart, Nicholas  
APPLICANT: Ricke, Darrell  
APPLICANT: Zhu, Tong  
TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION  
FILE REFERENCE: 60111-NP  
CURRENT APPLICATION NUMBER: US/10/260,238  
CURRENT FILING DATE: 2002-09-26  
PRIOR APPLICATION NUMBER: US 60/325,448  
PRIOR FILING DATE: 2001-09-26  
PRIOR APPLICATION NUMBER: US 60/325,277  
PRIOR FILING DATE: 2001-09-26  
PRIOR APPLICATION NUMBER: US 60/370,620  
PRIOR FILING DATE: 2002-04-04  
NUMBER OF SEQ ID NOS: 6077  
SEQ ID NO 565  
LENGTH: 1926  
TYPE: DNA  
ORGANISM: Oryza sativa  
FEATURE:  
NAME/KEY: N region  
LOCATION: (261)..(261)  
OTHER INFORMATION: n = any nucleotide  
FEATURE:  
NAME/KEY: N region  
LOCATION: (291)..(291)





PRIOR APPLICATION NUMBER: US 60/325,277  
PRIOR FILING DATE: 2001-09-26  
PRIOR APPLICATION NUMBER: US 60/370,620  
PRIOR FILING DATE: 2002-04-04  
NUMBER OF SEQ ID NOS: 6077  
SEQ ID NO 30  
LENGTH: 1759  
TYPE: DNA  
ORGANISM: Oryza sativa  
FEATURE:  
NAME/KEY: N region  
LOCATION: (1584)  
OTHER INFORMATION: n = any nucleotide  
FEATURE:  
NAME/KEY: N region  
LOCATION: (1587)..  
OTHER INFORMATION: n = any nucleotide  
FEATURE:  
NAME/KEY: N region  
LOCATION: (1602)..  
OTHER INFORMATION: n = any nucleotide  
US-10-260-238-30

Query Match 25.2%; Score 698; DB 15; Length 1759;  
Best Local Similarity 68.4%; Pred. No. 2.1e-160;  
Matches 1174; Conservative 0; Mismatches 438; Indels 105; Gaps 11;

```
Qy 171 ATGATACCGCGCTGGACCTTACAGTGTCTGACGGCGGTGGTGGCGCTGTACTGGCC 230
Db 1 ATGATAACGGGGCGGACTTCTACCGTGTATGATGCGCGGATGGTGGCGCTGTACGTGGG 60

Qy 231 ATGACGCTGGGTACGGCTCGCTCGGTGGTGGCGCATCTTACCGCGGACCAAGTGTCC 290
Db 61 ATGATCTGGGTACGGGTGGTGGCGCATCTTACCGCGGACCAAGTGTCTCC 120

Qy 291 GGGATCAACCGCTTGGTGGCGCTCTTGGCGGTGGCGCTCTCTCTCTCTCTCTCTCTCT 350
Db 121 GGGATCAACCGCTTGGTGGCGCTCTTGGCGGTGGCGCTCTCTCTCTCTCTCTCTCTCT 180

Qy 351 ACCAAGACCCCTTGGCGCATGACCGCTTCTGGCGCGGACACAGCTGCGAAGGTG 410
Db 181 ACCAAGACCCCTTACCATGACCTTGGCTTCTGGCGCGGACACAGCTTGGCGAGCTC 240

Qy 411 GCGGTCTTGGCGCTGCTGGCGCTGGCGCTTGGCGCGGCTCTCTCTCTCTCTCTCTCT 470
Db 241 ATGGTCTTGGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 291

Qy 471 CTCGACTGGAGATCACGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 530
Db 292 CTCGAGTGGACCATCACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 351

Qy 531 ATCCCGCTGCTGGAGGATGTAGCGCGCTGCTGGCGCGGACGCTCATGGTCCAGGTC 590
Db 352 ATCCCGCTGCTGAAGGATGTACGGGAGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 408

Qy 591 GTGGTCTTCCAGTGCATCATCTGTGTACAGCTGTCTCTCTCTCTCTCTCTCTCTCTCT 650
Db 409 GTGGTCTTCCAGTGCATCATCTGTGTACAGCTGTCTCTCTCTCTCTCTCTCTCTCTCT 468

Qy 651 GCGCGCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 710
Db 469 GCGAGATCTCTCATCACGAGATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 525

Qy 711 CCGCTGCACTCCAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 770
Db 526 GTGGTGGACCGCGCTGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 585

Qy 771 GACGGGTGCGCGCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 830
Db 586 GA-----GGTGAAGGAGGACCGGAGATAC 612

Qy 831 GTACCGTGGCAAGTTCACCAAGCTCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 887
```

```
Db 613 GTACCGTGGCGCGCTCCAAAGCGGTGCGCTCCGAGCTTACTCCGCGGCTCCATGGG 672
Qy 888 TCCAGACCATGACGCCCGCTGTGTCCAACTCTCCGCGTGGAGATCTACTCGTGGAG 947
Db 673 TTCTCCAGCACCGCCGCGCGGCGGAGCAACTCTACCAACGCGAGATCTACTCGTGGAG 732
Qy 948 TCGTGGCGCAACCCACCGCGCGGCTCCAGCTTCAACCAAGCGCGCTTCTTCAACATC 1007
Db 733 TCGTGGCGCAACCCACCGCGCGGCTCCAGCTTCAACCAAGCGCGCTTCTTCAACATC 792
Qy 1008 GTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1063
Db 793 GTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 852
Qy 1064 -----ATGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1107
Db 853 CCGCGCGCGCTCCAACTACGAGGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 912
Qy 1108 TCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1167
Db 913 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 972
Qy 1168 CCGTGTCCGAGCGCG-----CGCGGTGCGCGCTTCTGGCGCGCGCGCGCGCGCGCGCG 1218
Db 973 CCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1032
Qy 1219 ATGCCGAGCTCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1278
Db 1033 TGTCCGAGCTCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1092
Qy 1279 CGAGCAGGAGCAAGAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1338
Db 1093 GCGTGGAGGAGCGACTTTCAGCTTTCGCGGAAACAGAGGCGCTCGCGCGCGCGCGCGCG 1152
Qy 1339 TGGGTTCGAATCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1398
Db 1153 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1194
Qy 1399 CCGTGGCGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1458
Db 1195 CGAGCGCGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1254
Qy 1459 AGCTGATCGGAAACCCACACCTTCTCCAGCTTCTCCAGCGCTCATCGCGCGCTCTCTCT 1518
Db 1255 AGCTCATCGGAAACCCACACCTTCTCCAGCGCTCATCGCGCGCTCTCTCTCTCTCTCTCT 1314
Qy 1519 CCT---ACAGGTGGGCGATCGAGATGCCAGCGATCATCGCGCGCGCGCGCGCGCGCGCG 1575
Db 1315 GCTTCCGCGAGGTGGAACTTCGAGATGCCGCGCGATCATCTGAAATCCATCTCCATCTCT 1374
Qy 1576 CGGACCGCGGTCTCGGAGTGGCGATGTTTCCAGCT-AGGCTCTGTTTATGGCGCTCGAGCGCG 1634
Db 1375 CGGACCGCGGTCTCGGCGATGGCGATGTTTCCAGCTTCTCCAGCGCTTCTCCAGCGCGCT 1434
Qy 1635 AGGATCATCGGTGCGGGAACAAAGCTGGCGCGCATCGCGATGGGCGGTCCGCGTTCGTCGCA 1694
Db 1435 CGGATCATCGGTGCGGGAACAAAGCTGGCGCGCATCGCGATGGGCGGTTCCTCTGACCG 1494
Qy 1695 GCGCGCGCGGTCTATGGCGCGCGCGCTTCCATCGCGGTGCGGTCTCGCGCGCGCTCTCTCT 1754
Db 1495 GCGCGCGCGGTCTATGGCGCGCGCGCTTCCATCGCGGTGCGGTCTCGCGCGCGCTCTCTCT 1554
Qy 1755 ATGCCCATCGT---CCAGGCTGCTCTGCTCCAGGAGTCTGCGCGGTCTGTTTCCGCGAAG 1811
Db 1555 GTGCCCATGTGAGCGAGGAGCGGTGCGCGCGGTGCGCGGTGCGCGGTGCGCGGTGCGCGGT 1614
Qy 1812 GAGTACGGGTTCATCCCGACATCTTCAGCAGCGGT 1848
Db 1615 GAGTACAGGTGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1651
```

RESULT 3  
US-10-425-114-1349





Qy 1690 TCCAGGCCCCCGGCTCATGGCGCGCGCTTCCATCCCGCTCGGTTCGGCGCGCTCTCC 1749  
Db 1791 TCGCGCGCTGCGCTCATGGCGCGCGCTTCCATCCCGCTCGGTTCGGCGCGCTCTCC 1850  
Qy 1750 TCCACATCCGCTCATGGCGCGCGCTTCCATCCCGCTCGGTTCGGCGCGCTCTCC 1809  
Db 1851 TCGGATCGCCATGTGCGAGGCGCGCTTCCATCCCGCTCGGTTCGGCGCGCTCTCC 1910  
Qy 1810 AGGATACGCGCTTCCATCCCGCATCTTCCAGCAC 1843  
Db 1911 AAGATACACAGCTCCACCTCTGAGCAC 1944

RESULT 5  
US-10-260-238-5337  
; Sequence 5337, Application US/10260238  
; Publication No. US20040016025A1  
; GENERAL INFORMATION:  
; APPLICANT: Budworth, Paul R.  
; APPLICANT: Moughamer, Todd G.  
; APPLICANT: Briggs, Steven P.  
; APPLICANT: Cooper, Bret  
; APPLICANT: Glazebrook, Jane  
; APPLICANT: Goff, Stephen A.  
; APPLICANT: Katagiri, Fumiyaki  
; APPLICANT: Kreps, Joel  
; APPLICANT: Provart, Nicholas  
; APPLICANT: Ricke, Darrell  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION  
; FILE REFERENCE: 60111-NP  
; CURRENT APPLICATION NUMBER: US/10/260,238  
; CURRENT FILING DATE: 2002-09-26  
; PRIOR APPLICATION NUMBER: US 60/325,448  
; PRIOR FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: US 60/325,277  
; PRIOR FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: US 60/370,620  
; PRIOR FILING DATE: 2002-04-04  
; NUMBER OF SEQ ID NOS: 6077  
; SEQ ID NO 5337  
; TYPE: DNA  
; ORGANISM: Zea mays  
US-10-260-238-5337

Query Match 21.2%; Score 586.4; DB 15; Length 651;  
Best Local Similarity 99.7%; Pred. No. 2.8e-133;  
Matches 598; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 1238 AGGAGCCAGGCTTACGACGAGTACGGGCGCGCGCGCTACGACGAGCAGGACGAG 1297  
Db 52 AGGAGCCAGGCTTACGACGAGTACGGGCGCGCGCGCTACGACGAGCAGGACGAG 111  
Qy 1298 GAGCGGCGCGCGCGCGCGCGCGCGCGCGCTTCCATCCCGCTCGGTTCGGCGCGC 1357  
Db 112 GAGCGGCGCGCGCGCGCGCGCGCGCGCGCTTCCATCCCGCTCGGTTCGGCGCGC 171  
Qy 1358 GCAGCTGTACCCCAAGGACGACGCGGAGGAGGCGCGCGCGCTTCCATCCCGCTCG 1417  
Db 172 GCAGCTGTACCCCAAGGACGACGCGGAGGAGGCGCGCGCGCTTCCATCCCGCTCG 231  
Qy 1418 GAGCGTATGACGCGGCTCATCTTCCATGCTGTGAGGAGCTGATCCGGAACCCCA 1477  
Db 232 GAGCGTATGACGCGGCTCATCTTCCATGCTGTGAGGAGCTGATCCGGAACCCCA 291  
Qy 1478 CACCTACTCCAGCCTCATCGCGCTCGTCTGCTTCCCTTCCATGAGTGGGCGCATCGA 1537  
Db 292 CACCTACTCCAGCCTCATCGCGCTCGTCTGCTTCCCTTCCATGAGTGGGCGCATCGA 351  
Qy 1538 GATGCGAGCATCATCGCGCGCTCGATTCGATCCTGTCCGACGCGGCTTCGGGATGCG 1597

Db 352 GATGCCAGCGATCATCGCCCGGTCGATTCGATCCCTGTTCGACGCGGCTTCGGGATGCC 411  
Qy 1598 CATGTTTCAGCTAGGCTGTTCATGCGCTGACGCGAGGATCATCGCTGCGGGAACAA 1657  
Db 412 CATGTTTCAGCTAGGCTGTTCATGCGCTGACGCGAGGATCATCGCTGCGGGAACAA 471  
Qy 1658 GCTGGCGGCGCATCGCATGCGGCTTCGCTTCGAGGCGCGCG-GGTTCATGCGCGCGC 1716  
Db 472 GCTGGCGGCGCATCGCATGCGGCTTCGCTTCGAGGCGCGCGCGGTCATGCGCGCGC 531  
Qy 1717 CTTCCATCGCGCTGCTTCGCGGCGCTTCTTCCATCGCATGCTTCAGGCTGCTC 1776  
Db 532 CTTCCATCGCGCTGCTTCGCGGCGCTTCTTCCATCGCATGCTTCAGGCTGCTC 591  
Qy 1777 TGCCTCAGGCGATCGCTCGCTTCGCTTCGCGAGGAGTACGCGTTTCATCCGACATCC 1836  
Db 592 TGCCTCAGGCGATCGCTCGCTTCGCTTCGCGAGGAGTACGCGTTTCATCCGACATCC 651

RESULT 6  
US-10-424-599-134028  
; Sequence 134028, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic, David K  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 134028  
; LENGTH: 2346  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_92032C.1  
US-10-424-599-134028

Query Match 16.3%; Score 451.2; DB 12; Length 2346;  
Best Local Similarity 58.1%; Pred. No. 6.3e-100;  
Matches 1020; Conservative 0; Mismatches 643; Indels 93; Gaps 9;

Qy 164 CGCGCCATGATCACCGCGCTTGACCTTACACGCTGCTGACGCGGCTGCTGCGCTGTA 223  
Db 141 CACGACATGATAACGCTTGTGACCTTGTACACGCTTGTGACGCGGCTGCTGCGCTTA 200  
Qy 224 CGTGGCATGACGCTGCGGTACGCTCGCTCGGTGCGGCTTCTTACGCGGACCA 283  
Db 201 CGTGGCATGACCTTCCGCTACGCTCGGTGAGTGTGAGATCTTACCGCGGACCA 260  
Qy 284 GTGCTCGGGATCAACCGCTTTCGCGCTTTCGCGCTGCGCTTCTTCTTCCACTT 343  
Db 261 ATGCTCGGGAATAACCGCTTTCGCTGCGCTTTCGCGCTTCTTCTTCCACTT 320  
Qy 344 CATCTCCACCAACGCGCTTTCGCGCTTTCGCGCTTTCGCGCTTTCGCGCTTTCGCG 403  
Db 321 CATCTCCACCAACGCGCTTTCGCGCTTTCGCGCTTTCGCGCTTTCGCGCTTTCGCG 380  
Qy 404 GAAGTGGCGCTTTCGCGCTTTCGCGCTTTCGCGCTTTCGCGCTTTCGCGCTTTCGCG 463  
Db 381 AAAGCCATGATCTTCGCTGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCT 431  
Qy 464 GCTCGGCTTCGCTGAGCATGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCT 523  
Db 432 AGGCTTCATGAGTGTTCATCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 491  
Qy 524 CATGGGCTTCCGCTGCTGCGGAGCATGATACGCGCGCTTCGCGCGGCTTCGCTGCTG 583  
Db 492 GATGGGCTTCCGCTGCTGAGGAGCATGATACGCTGATGACTCG---GGAACTAATGGT 548

QY 584 CAGGTGCTGCTCTCCAGTGCATCATCTGTGTAACAGCTCANGCTCTCTCTTCGAGTA 643  
Db 549 TCAGATTGCTGCTTCCAGTGTATCATCTGTATACCTTATGTGTTCTTTTCGAGTA 608  
QY 644 CCGGCGCGCGCGCGCTCTCGACAGTTCCTCCGAGCGCGCCCGGCTCCATCGT 703  
Db 609 TAGGGTGGAGGGCTTTAAATAGTGGAGCAGTTTCCTGA---TACAGGGGTTCGATAT 665  
QY 704 CTCCTTCGCTCGCATCCGACGTGCTCTCGCTCGCAGGGGGGACGTGAGCTCGAGGC 763  
Db 666 CTCCTTCAAGTTGATTCCGATATTATTTCTGTGGTGAAGAGCGCTTCAGACGA 725  
QY 764 CGAGCCGACGGC-----GTCCGCGCGCGCGC 791  
Db 726 GCGGAGGTTGTTGATGACGGGAAGCTTCATGTGACGGTGAAGTCCGGAGTTCCGG 785  
QY 792 GCCGTCTCTCCCGCGCGGGAGCGCCGGGGGGTGGCGGTCACTCCAGACCCCGC 847  
Db 786 GTCGGAGATTTTCTCTAGCGCTCTCATGGCCCCCAACTCGGTGCTGTTGACTCCGAGGCC 845  
QY 848 CACAGCTCGGCTCCGAGGCGGGTGTCTCGCATCTCGCATCCAGACCATCGACGCCG 907  
Db 846 TTCGAATTTGACCAATCGGAGATTTACTCGCTCGAGCTCGAGGAACCGACTCCGAG 905  
QY 908 TGTGTCCAACTCTCCGCGTGGAGATCTACTCGCTGCAGTCTGCGGCAACCCACCCC 967  
Db 906 AGGCTCGATTTTAAACACACGAGTTTTTATCTCATGTTGATGTTGAACACACAA 965  
QY 968 GCGGGTCCAGCTTCAACACGCGGACTTCTTCAACATGCTCGGCGCGCGCCCAAGG 1027  
Db 966 CAACAAACAACACGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1025  
QY 1028 AGCGGAGGAGCGCGGGGACGAGG-----AGAGGGCGCATCGCGCGC 1073  
Db 1026 AAGTGTGGGGTGTGAGGTTGAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1085  
QY 1074 GCGGCGGAGGACACT-----CGCGGAGCGCGGCGCGCTGCGCGTGCCTGCGGCA 1123  
Db 1086 GAATCGGGGATTTTCTCGGTGGCGCGGAGAGAGGTTGTGAGAGTGTGTGTGTGT 1145  
QY 1124 GAGGAAGACCTGCATGCTGCTGAGCTCCAGCGCTCGCGCTCGGTCGAGCGCGC 1183  
Db 1146 AGGGAAGATCTCCACATGTTGTGTGAGTTGAGTGTCTCGCGGTGTCGGAAGTGG 1205  
QY 1184 CGCGTGCACGTCTTTCGCGCGCGCGCTGACCATGCGGAGCTCTCGCCAAAGAGC 1243  
Db 1206 AATCCATGCTTTCAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1265  
QY 1244 CCAGGCC-----TAGCAGAGTACGGGCGCGGACGACTACAGCAGCAGACGA 1291  
Db 1266 GGTGCTCACCAGAAAGATTATGATGATGATGATGATGATGATGATGATGATGATG 1325  
QY 1292 GAAAGGAGCGCGCGCGGACAGGGCGCGCGAGCTGTGGAAGCTGGGGTCCAACTC 1351  
Db 1326 AACCTTCTAATGGGTTGACAGGAAGGCGAGTGTCTTCAAGCTTGGCTCGAGTTC 1385  
QY 1352 GACGCGGAGCTGTATCCCAAGGACGAGCGGAGGGGAGGGCGCGGCTGCGGAGTCC 1411  
Db 1386 CACAGTGAAGTTCGCCCTTAAAGCTCAAGGTGAAGCCAAACCTAC-----TTCCATGCC 1439  
QY 1412 GCGGCGGAGCTGATGACGCGCTCATCTCATATGTTGTGAGGAAGCTGATCCGAA 1471  
Db 1440 ACCAAAGAGTTATGACAGGCTCATTTGATGATGATGATGATGATGATGATGATGAT 1499  
QY 1472 CCCCACACCTACTCCAGCTCATCGGCTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1531  
Db 1500 CCCCACACATATCCAGCTCTTTGGTCTCACTTGTCTTGTATCTCAITCAATCAATGAA 1559  
QY 1532 CATGAGATGCGAGCATCATCGCGGTGAGTTTCGATCTGTGCGAGCGGGTCTCGG 1591  
Db 1560 TGTGTTATGCGAGCAATTTGTTGCTAAATCGATATCAATTTTATGATGATGCTGTGG 1619

QY 1592 GATGGCATTTTCAGCTAGGCTGTTTATGCGGCTGCGAGCGAGGATCATCGCGTGGC 1651  
Db 1620 GATGGCAATTTTTCAGCTTGGCTTATTCATGCGATTCGAGCCAAAGATTTTTCATGTGG 1679  
QY 1652 GAACAGCTGGCGGCGCATCGCATGGGCGTCCGTTCTGTCGAGGCGCCCGCGTTCATGCG 1711  
Db 1680 AACTCGGTTGCTTCTTTGCTATGCGAGTTCGTTTCTTACTGCTCTGCACTCATGCG 1739  
QY 1712 CGCGGCTTCATCGCGCTCGGCTCGGCTCGGCGGCGCTCTCTCCACATCGCATCGTCCAGC 1771  
Db 1740 TGTGCTTCAATCGTTGTAGGCTCAGGGAGTTCTATTGCAATGCTTATTGTACAGC 1799  
QY 1772 TGTCTGCTCAGGGGATCGTCCGCTTCTGTCGCAAGGAGTACGGGTTTCATCCCCA 1831  
Db 1800 TGTCTGCTCAGGGGATGTTCCCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1859  
QY 1832 CATCTGAGCAGCG 1847  
Db 1860 CATACTAAGCACCGG 1875

RESULT 7  
US-10-425-114-15739  
; Sequence 15739, Application US/10425114  
; Publication No. US2004003488A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaka, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 15739  
; LENGTH: 1273  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3060-064-E7\_F1J  
US-10-425-114-15739

Query Match 16.3%; Score 450; DB 12; Length 1273;  
Best Local Similarity 66.7%; Pred. No. 9.4e-100;  
Matches 743; Conservative 0; Mismatches 320; Indels 51; Gaps 5;

QY 171 ATGATCACCGCGTGGACCTCTACACGCTGCTGACGCGGCTGCTGCGCTGTACGTGGC 230  
Db 188 ATGATTCGCGCGCGGACTTGTACACGATGATGACGCGCTGTGCGCTGTACGTGGC 247  
QY 231 ATGACGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 290  
Db 248 ATGATCCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 307  
QY 291 GGGATCAACGCTTCTGCGGCTCTTTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 350  
Db 308 GGGATCAACGCTTCTGCGGCTCTTTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 367  
QY 351 ACCAAGACCCCTTCCCATGAACCTTCTGCGCTGCTGCGCGCGGACACGCTGCAAGAGTG 410  
Db 368 ACCAAGACCCCTTACACCATGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 427  
QY 411 GCGCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 467  
Db 428 ATGCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 487  
QY 468 GGGCTGCTGAGCATCGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 527  
Db 488 GACCTGCTGAGCATCGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 547



```
QY 528 GGCATCCCGTCTGTCAGGACATGTACGGCGGTGTCGGCGGACGCTCATGTGTCAG 587
Db 548 GGCATCCCGTCTGTCAGGCGCATGTACGGCGACTTCTCG---GGAAGCCTCATGTGTCAG 604
QY 588 GTCTGCTCTCTCGAGTGCATCTGTGTACAGCTCATGCTCTTCTCTTCTGAGTACGGC 647
Db 605 ATCGTCTGCTCTCGAGTGCATCTGTGTACAGCTCATGCTCTTCTGAGTACGGC 664
QY 648 GCGCGCGCGCTCTGCTCTCGACCATGTTCCCGGAGCGCGCGCTTCCATGCTCTCC 707
Db 665 GCGCGCGCTCTCTCATCGCGGAGAGTTCCCGGCAACGCGCGCGCTCATGCTCTCC 724
QY 708 TTCGCGTCTGATTCGAGCTGTCTGTCTGTCGCGAGGGGGGAGCTGAGCTGAGGCCGAG 767
Db 725 ATCGCGCTGACCGCGAGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 757
QY 768 CCGAGCGGGTCTCGCGCGCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 827
Db 758 GACACGGCGCGGAGCGCATCGAGAGCGAGCGCGAGGTTCAGGGAGAGCGGAGGATA 817
QY 828 CCGCTCACTGCGTGGGCAAGTTCACAGCTCGCGTTCGAGCGCGCTGCTGCGAC----- 881
Db 818 CAGCTCAGCTGCGCGCTCTCAAGCGCTGCGCTCTGAGCTCTCTCTCTCTCTCTCTCTCTCTCT 877
QY 882 ---TCGACTCCAGACCATGACAGCGCGCTGTGTCTCAAGCTCTCTCGGGGTGAGATCTAC 938
Db 878 GTGGCGGCTTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 937
QY 939 TCCTGAGTCTGCGCGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 998
Db 938 TCCTGAGTCTGCGCGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 997
QY 999 TTCAATCATGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1058
Db 998 TACTCATGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1057
QY 1059 GCGCGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1109
Db 1058 ACGCCGCGCGCGTCCAACTACGAGGAGCGCGTCCAGCGCGCGCGCGCGCGCGCGCGCGCG 1117
QY 1110 GCGGTGCGCGCGCAAGAGGAGGACCTGTGACATGCTGTGAGCTTCCAGCGCGCTCGGCC 1169
Db 1118 GTGGCAGGCGCACTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1177
QY 1170 GTCTCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1229
Db 1178 AAGGCGAGGACCTCCACATGTTCTGTCTGAGCTCCAGCGCTCCAGCGCGCTCCAGCGCTC 1237
QY 1230 CTCGCCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1263
Db 1238 TTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1271
```

## RESULT 8

```
US-10-425-114-11838
; Sequence 11838, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 11838
; LENGTH: 1313
```

```
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 701184321_FLI
US-10-425-114-11838

Query Match 15.4%; Score 427.8; DB 12; Length 1313;
Best Local Similarity 72.0%; Pred. NO. 2.6e-94;
Matches 650; Conservative 0; Mismatches 202; Indels 51; Gaps 5;

QY 164 CCGCGCATGATCACCGCGCTGAGACCTCTACCACTGCTGACGCGCGGTGTCGCTGTA 223
Db 188 CTCACGATGATTACGGGACGACTTCTACACGCTCATGACGCGGTGTCGCTGTA 247
QY 224 CTTGGCATACGCTGCGTACGCTCCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 283
Db 248 CTTGGCATATCTGCGGTACGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 307
QY 284 GTCTCCGGGATCAACCGCTTCTGCGCTCTTTCGCGTCTCTCTCTCTCTCTCTCTCTCTCT 343
Db 308 GTCTCCGGGATCAACCGCTTCTGCGCTCTTTCGCGTCTCTCTCTCTCTCTCTCTCTCTCT 367
QY 344 CATCTCCACCAACGACCGCTTTCGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 403
Db 368 CATCTCCACCAACGACCGCTTTCGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 427
QY 404 GAAGGTGCGCTCTCTGCGCTGCTGCGCTGCTGCGCTCTCTCTCTCTCTCTCTCTCTCTCT 463
Db 428 GAAG-----CTCATGCTGCTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 478
QY 464 GCTCGGCTCGATGAGGACATCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 523
Db 479 GGGCAGCTCGAGTGGACCATCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 536
QY 524 CATGGGATCCCGTCTGCGAGCATGTACGGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 583
Db 539 CATGGGATCCCGTCTGCTCAAGGCGATGTACGGGCTCTCTCTCTCTCTCTCTCTCTCTCT 595
QY 584 CCAGTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 643
Db 596 GCAGTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 655
QY 644 CCGCGCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 703
Db 656 CCGCGCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 712
QY 704 CTCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 763
Db 713 CTCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 772
QY 764 CGAGCGCGAGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 823
Db 773 GAGGCGCA-----GGTCAAGGAGGAGCGCAG 799
QY 824 GGTGCGCTCACCGTGGCAAGTCCACAGTCTGCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 880
Db 800 GATACAGTCACTGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 859
QY 881 CTCGACTCTCCAGACATGACGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 940
Db 860 CATGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 919
QY 941 GCTCAGTCTGCGCAACCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1000
Db 920 GCTCAGTCTGCGCAACCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 979
QY 1001 CAACATCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1060
Db 980 CTCATGCTGCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1039
QY 1061 CGC 1063
Db 1040 CGC 1042
```



## RESULT 9

US-10-425-114-2319  
; Sequence 2319, Application US/10425114  
; Publication No. US2004003488A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 2319  
; SEQ ID NO 2319  
; LENGTH: 1334  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700209522\_FLI  
US-10-425-114-2319

Query Match 15.4%; Score 427.8; DB 12; Length 1334;  
Best Local Similarity 72.0%; Pred. No. 2.6e-94;  
Matches 650; Conservative 0; Mismatches 202; Indels 51; Gaps 5;  
QY 164 CGCGCCATGATCAACCGCGTGGACCTTACACAGTGTGACGGGGTGGTGGCGCTGTA 223  
DB 210 CTCACAGTATACCGGGACGGACTTCTACACGTCATGACGCGCGTGGTGGCGCTGTA 269  
QY 224 CGTGGCCATGACGTGGCGGTACGCTCGCTCGGTGGGGCATCTTACGCGCGACCA 283  
DB 270 CGTGGCGATGATCTTGGCGTACGGTTCGGTGGTGGCGCATCTTTCGCGCGACCA 329  
QY 284 GTGCTCCGGGATCAACCGCTTGGTGGCGCTTTCGCGGTGGCGCTCTCTCTTCCACTT 343  
DB 330 GTGCTCCGGGATCAACCGCTTGGTGGCGCTTTCGCGGTGGCGCTCTCTCTTCCACTT 389  
QY 344 CATCTCCACCAAGACCGCTTGGCATCAACCTGCTTCTGCGCGCGGACACGCTGCA 403  
DB 390 CATCTCCACCAAGACCGCTTACACCATGCTGGCTTCTGCGCGCGGACACGCTGCA 449  
QY 404 GAAGTGGCGGCTCTCGCGCTGCTGGCGCTTCTCGCGCGCTTCTCTCTCGCGCGC 463  
DB 450 GAAG-----CTCATGGTGTGCGCATGCTCACGCGTGGAGCCACCTCAGCGCGC 500  
QY 464 GCTCGGGCTGACTGGAGATCAGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 523  
DB 501 GGGCAGCTGGAGTGGACCATCACGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 560  
QY 524 CATGGGCATCCCGCTGCTCGAGGATGTACGGCGCTGCTCGCGCGGACGCTCATGCT 583  
DB 561 CATGGGCATCCCGCTGCT 617  
QY 584 CCAGTGTGTGTCT 643  
DB 618 GCAGATGTGTGTCT 677  
QY 644 CGCGCGCGCGCGCGCTGCT 703  
DB 678 CGCGCGCGCGCGCGATGCTCATCAGGAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 734  
QY 704 CTCCTTCCGCGTGAATCCGATCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 763  
DB 735 CTCCTTCTGCTGACACCGGAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 794  
QY 764 CGAGCCGACGGGCTGCGCGCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 823

DB 795 GGAGGCCGA-----GGTCAAGGAGGACGGCAG 821  
QY 824 GGTGCGCGTCAACCGTGGCAAGTCCACAGCTCGCGTCCGA---GGCGCGGTGCTCGCA 880  
DB 822 GATACACGTACCGTGGCGCGCTCAACGCTTCGCGCTCCGACATCTACTCGCGCGCTC 881  
QY 881 CTCGCATCCACAGACATGACAGCCCGGTGTGTCAACCTCTCCGCGGTGGAGATCTACTC 940  
DB 882 CATGGGCTTCTCCAGCAACCGCGCGCGCGAGCAACCTGACCAACGCGCGAGATCTACTC 941  
QY 941 GCTGAGTGTGCGCGCAACCGCGCGCGCGCGGTTCAGCTTCAACCAACGCGCGCTTCTT 1000  
DB 942 GCTGAGTGTGCGCGCAACCGCGCGCGCGCGGTTCAGCTTCAACCAACGCGCGCTTCTT 1001  
QY 1001 CAACATCTGTCGCGCGCGCGCGCAAGGAGGCGCGCGGAGCGCGCGGAGGAGGAGG 1060  
DB 1002 TCCCATGTCGCGCGCGCGCTCAACTTTCGCGCGCGCGCGCTTCGCGCATCCGACCGG 1061  
QY 1061 CGC 1063  
DB 1062 CGC 1064

## RESULT 10

US-10-425-114-29155  
; Sequence 29155, Application US/10425114  
; Publication No. US2004003488A1  
; GENERAL INFORMATION:  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 29155  
; LENGTH: 1199  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB83-006-E4\_FLI  
US-10-425-114-29155

Query Match 15.4%; Score 426.2; DB 12; Length 1199;  
Best Local Similarity 71.9%; Pred. No. 6.1e-94;  
Matches 649; Conservative 0; Mismatches 203; Indels 51; Gaps 5;  
QY 164 CGCGGCCATGATCACCGCGCTGGACCTCTACACAGTGTGACGGGGTGGTGGCGCTGTA 223  
DB 77 CTCACAGTATACCGGGACGGACTTCTTACACGTCATGACGGCGTGGTGGCGCTGTA 136  
QY 224 CGTGGCCATGACGTGGCGGTACGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 283  
DB 137 CGTGGCGATGATCTCTGCGCTACGGGTGGTGGCGGTGGTGGCGCATCTTCTCGCGGACCA 196  
QY 284 GTGCTCGGGATCAACCGCTTCTGTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 343  
DB 197 GTGCTCGGGATCAACCGCTTCTGTCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 256  
QY 344 CATCTCCACCAAGACCGCTTCTGCGATGAACCTGCGCTTCTGCGCGCGGACGACGCTGCA 403  
DB 257 CATCTCCACCAAGACCGCTTACACCATGAACCTGCGCTTCTGCGCGCGGACGACGCTGCA 316  
QY 404 GAAGTGGCGGCTCTCGCGCTGTCGCGCTTCTGCGCGCGGCTCTCTCTCTCTCTCTCTCTCT 463  
DB 317 GAAG-----CTCATGGTGTGCGCATGCTCACGGGTGGAGCACTCTCAGCGCGCG 367  
QY 464 GCTCGGCTCGACTGGAGCATCAGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 523

```

Db 368 GGGCAGCCTGGAGTGGACATCACGCTCTTCTCCCTCTCCAGCTGCGCCACACAGCTCGT 427
Qy
Qy 524 CATGGGCATCCCGCTGCTGCGAGGATGTACGGCGGTCGTCGCCCGGCGAGCTCATGGT 583
Db 428 CATGGGCATCCCGCTGCTCAAGGGCATGTACGGCGACTTCT---CCGGCAGCCTCATGGT 484
Qy 584 CGAGTCTGCTGCTCCAGTGCATCATCTGGTACACGCTCATGCTCTTCTCTTCGAGTA 643
Db 485 CGAGTCTGCTGCTCCAGTGCATCATCTGGTACACGCTCATGCTCTTCTTCGAGTA 544
Qy 644 CGCGCCGCGCGCGCTGCTCTCCAGCAGTTCGCCGAGCGCGCGCGCTCCATCGT 703
Db 545 CGCGCGCGCGGATGCTCATCACGAGCAGTTCCCGGA---CAAGCGCGGGGCGCATCGC 601
Qy 704 CTCCTTCGCGTGCAGTCTCGAGTCTGCTCGTCCGAGGGGAGCTCGAGCTCGAGGC 763
Db 602 CTCCTTCGCTGCGACCGGAGCTGCTCTCCGCTCGACCGCGCGGAGCGCATCGAGAC 661
Qy 764 CGAGCCGAGCGCTGCGCGCGCGCGCGCTCTCTCCGCGCGCGCGGCGAGCGCGGCG 823
Db 662 GAGGCGGA-----GGTCAAGGAGGAGCGGCGAG 688
Qy 824 GGTGCGGTCACGTCGCGAAGTCCACAGCTCGCGCTCCGAGCGCGGCTGCTCGCACTC 883
Db 689 GATACACGTCACCGTGGCGCGCTCCACGCGCTCGCGCTCCGACATCTACTCCCGCGCTC 748
Qy 884 ---GCACTCCAGACCATCAGCCCGCTGTGTCCAACTCTCCCGCGCTGAGATCTACTC 940
Db 749 CATGGCTTCTCCAGCACACCGCGCGCGCGCGAGCACTGACCAACGCGGAGATCTACTC 808
Qy 941 GTGAGTCTGTCGCGCAACCCACCGCGCGCGGTCGAGTTCACACGCGCGACTTCTT 1000
Db 809 GCTGAGTCTGTCGCGCAACCCACCGCGCGCGGCTCCAGTTCACACGAGCTTCTTA 868
Qy 1001 CAACATCTGTCGCGCGCGCGCGCGAGGAGCGGCGGAGCGCGGCGGAGGAGGAGG 1060
Db 869 CTCCTATGTCGCGCGCGAGTCCAACTTCTCGCGCGCGCGAGCGGCTCGGATCCGACCGG 928
Qy 1061 CGC 1063
Db 929 CGC 931

```

```

RESULT 11
US-10-425-114-33502
; Sequence 33502, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 33502
; LENGTH: 1321
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17127C09_FLI
US-10-425-114-33502

```

```

Query Match 15.4%; Score 425.6; DB 12; Length 1321;
Best Local Similarity 72.1%; Pred. No. 9e-94;
Matches 646; Conservative 0; Mismatches 199; Indels 51; Gaps 5;

```

```

Qy 171 ATGATCACCGCGCTGGACCTCTACACGTCGTGCTGAGCGGGTGGTCCCGTGTACGTGGCC 230
Db 215 ATGATTCAGGGGACGAGCTTCTACACGTCATGACGGCGGTGTGCGGTGTACGTGGCG 274
Qy 231 ATGACGCTGGCGTAAGCTCCGTCGCTGGTGGCGCATCTTACAGCGCGGACCGAGTGTCC 290
Db 275 ATGATCTCGGCTACGGGTGCGGTGGTGGCGCATCTTCTCGCGGACCGAGTGTCTCC 334
Qy 291 GGGATCAACCGCTTCTGGTGGCTCTTTCGCGCTGCGCTCTCTCTCTCTCTCTCTCTCTCT 350
Db 335 GGGATCAACCGCTTCTGGTGGCTCTTTCGCGCTGCGCTCTCTCTCTCTCTCTCTCTCTCT 394
Qy 351 ACCAAGCAACCGCTTTCGCGCATGAACTGGCTTCTTGGCGCGCGACACGCTGCAAGGTG 410
Db 395 ACCAAGCAACCGCTTTCGCGCATGAACTGGCTTCTTTCGCGCGCGACACGCTGCAAGG 451
Qy 411 GCGCTCTCTCGCTGCTGGCTGCGCTTCCCGGCGCTCTCTCTCTCTCTCTCTCTCTCTCT 470
Db 452 -----CTATGTTGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 505
Qy 471 CTCGACTGGAGCATACGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 530
Db 506 CTGGAGTGGACCATACGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 565
Qy 531 ATCCGCTCTGCGAGGATGACGCGCTGCTGCGCGCGCGCGCTCTCTCTCTCTCTCTCTCT 590
Db 566 ATCCGCTCTGCTCAAGGGCATGACGCGCTTCT-----CCGGCAGCTCTCTCTCTCTCT 622
Qy 591 GTGCTCTCTCAGTGCATCATCTGGTACACGCTCATGCTCTTCTCTCTCTCTCTCTCTCTCT 650
Db 623 GTGCTCTCTCAGTGCATCATCTGGTACACGCTCATGCTCTTCTCTCTCTCTCTCTCTCTCT 682
Qy 651 GCGCGCGCTCTGCTCTCTGACGAGTTCGCCGCGCGCGCGCGCTCTCTCTCTCTCTCTCTCT 710
Db 683 GCGCGGATGCTCATACCGAGGAGTTCGCCGGA---CAACCGCGCGCGCTCTCTCTCTCTCT 739
Qy 711 CGGCTCTGACTTCCAGCTCTGCTCTGCTCGCGAGGGGAGCTGAGTCTGAGGCGCGAGCGCC 770
Db 740 GTGCTGACCGGAGCTGCTCTCTCTGACGCGCGCGAGGAGCGCCATCGAGACGAGGCG 799
Qy 771 GAGCGGTCTCCCGCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 830
Db 800 GA-----GGTCAAGGAGGACGGCAGATACAC 826
Qy 831 GTCACTGTCGCAAGTCCACAGCTCTCGCTCCGA---GGCGCGCTGCTGCTCTCTCTCTCT 887
Db 827 GTCACTGTCGCGCTCTCAACGCTCTCGCTCTCGACATCTTCTCTCTCTCTCTCTCTCTCT 886
Qy 888 TCCAGACCATGACGCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 947
Db 887 TTTCTCCAGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 946
Qy 948 TCGTCGCGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1007
Db 947 TCGTCGCGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1006
Qy 1008 GTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1063
Db 1007 GTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1062

```

## RESULT 12

```

US-10-425-114-3558
; Sequence 3558, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

```

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 3558

; LENGTH: 1091

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: 700321691\_FLI

US-10-425-114-3558

Query Match 14.9%; Score 412.8; DB 12; Length 1091;  
Best Local Similarity 71.2%; Pred. No. 1.1e-90;  
Matches 638; Conservative 0; Mismatches 207; Indels 51; Gaps 5;

Qy	171	ATGATCACCGGCTGGACCTTACACGCTGCTGACGGCGGTGGTGGCGCTGTAGTGGCC	230
Db	12	ATGATCACCGGACGGACTTCTACACGCTATGACGGCCATGGTGGCTGTAGTGGCC	71
Qy	231	ATGAGCTGGGCTAGCGTCCGTCGGTGGTGGCGCATCTTCACGCGGACAGTGTCC	290
Db	72	ATGATCTCGGCTACGGATCGCTCAGGTGGTGGCGCATCTTCACGCGGACAGTGTCC	131
Qy	291	GGGATCAACCGCTTCGTGGCGCTCTTCGCGTGGCGCTCTCTCTCTTCCACTTCATCTCC	350
Db	132	GGGATCAACCGCTTCGTGGCGCTCTTCGCGTGGCGCTCTCTCTCTTCCACTTCATCTCC	191
Qy	351	ACCAACGACCCCTTCGCATGAACCTGCGCTTCCTGGCGCGGACACGCTGCGAAGTG	410
Db	192	ACCAACGACCCCTTACCATGAACATCGGTTTCATCGCGCGGACACGCTGCGAAGTG	251
Qy	411	GCCTGCTTCGCGCTGCTGGCGCTGCGCTCCGCGCGCTCTCTCTCCCGCGCGCTCGGG	470
Db	252	ATCGTCTTCGCGCTGCT-----CACTGCATGGAGTACTCTCTCCGCGGGTCTC	302
Qy	471	CTCGACTGGACATACGCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	530
Db	303	CTCGAGTGGACATACGCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	362
Qy	531	ATCCGCTGCTCGAGGATGATGAGGCGCTGCTGGCGCGCTCTCTCTCTCTCTCTCTCTCT	590
Db	363	ATCCGCTGCTCGAGGATGATGAGGCGCTGCTGGCGCGCTCTCTCTCTCTCTCTCTCTCT	419
Qy	591	GTCGCTCTCCAGTGCATCTGTGTACACGCTCATGCTCTCTCTCTCTCTCTCTCTCTCTCT	650
Db	420	GTGGTCTCTCCAGTGCATCTGTGTACACGCTCATGCTCTCTCTCTCTCTCTCTCTCTCT	479
Qy	651	GGCGCGGCTGCTCTCGACGAGTTCCTCGACGGCGCGCGCTGCTGCTCTCTCTCTCTCT	710
Db	480	GCCAGGA*CTCTATCACCAGCAGTTCCTCCGACACGGCGGCG---CCATCGCTCTCATC	536
Qy	711	CGCGTCGACTCCGAGCTGCTCTCGCTCGCCAGGGGGACGCTGAGCTCGAGCGCGAGCCC	770
Db	537	GTGGTGGACCCCGAGCTGCTGCTGTGAACGGGCGCAACGACGCTGAGACGAGAGGCC	596
Qy	771	GACGCGCTCGCGCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	830
Db	597	GA-----GGTGAAGAGGAGCAGCAAGATACAC	623
Qy	831	GTCACCGTGGCAAGTCCACAGCTCGCGCTCCGAGGCGCGCTCTCTCTCTCTCTCTCTCT	887
Db	624	GTACCGTGGCGGCTCTCAACCGCTCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	683
Qy	888	TCCAGACCATGACCCCGCTGTGTCTCAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	947
Db	684	TTCTCCAGCACACCGCGCGCGCGCAACCTGACCAACGCGCGAGATCTACTCTGCTGAG	743
Qy	948	TGCTCGGAAACCCACCCCGCGCGGCTCGAGTTCACACGCGGAGTCTTCAATC	1007
Db	744	TGCTCGAGAAACCCACCGCGCGGCTCTCAACCTCTCTCTCTCTCTCTCTCTCTCTCTCT	803

Qy	1008	GTCGCGCGCGCGCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1063
Db	804	GTCGCGCGCGCAGCTCCAACTTCGCCCGCGGAGCGGCTTCGGCTGCGCAGCGGCGC	859

RESULT 13

US-10-260-238-5464/c

; Sequence 5464, Application US/10260238

; Publication No. US2004016025A1

; GENERAL INFORMATION:

; APPLICANT: Moughamer, Paul R.

; APPLICANT: Briggs, Steven P.

; APPLICANT: Cooper, Bret

; APPLICANT: Glazebrook, Jane

; APPLICANT: Goff, Stephen A.

; APPLICANT: Katagiri, Fumiaki

; APPLICANT: Kreps, Joel

; APPLICANT: Provart, Nicholas

; APPLICANT: Ricke, Darrell

; APPLICANT: Zhu, Tong

; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION

; FILE REFERENCE: 60111-NP

; CURRENT APPLICATION NUMBER: US/10/260,238

; CURRENT FILING DATE: 2002-09-26

; PRIOR APPLICATION NUMBER: US 60/325,448

; PRIOR FILING DATE: 2001-09-26

; PRIOR APPLICATION NUMBER: US 60/325,277

; PRIOR FILING DATE: 2001-09-26

; PRIOR APPLICATION NUMBER: US 60/370,620

; PRIOR FILING DATE: 2002-04-04

; NUMBER OF SEQ ID NOS: 6077

; SEQ ID NO 5464

; LENGTH: 702

; TYPE: DNA

; ORGANISM: Zea mays

US-10-260-238-5464

Query Match 13.4%; Score 370; DB 15; Length 702;

Best Local Similarity 78.4%; Pred. No. 2.7e-80;

Matches 486; Conservative 0; Mismatches 120; Indels 14; Gaps 3;

Qy	171	ATGATCACCGCGCTGGACCTTACACGCTGCTGACGGCGGTGGTGGCGCTGTACCTGGCC	230
Db	651	ATGATTCAGGGAGCGACTTCTACACGCTCATGACGGCGGTGGTGGCGCTGTACCTGGCC	592
Qy	231	ATGACGCTGGCGTACGGCTCCGCTCGCTGGTGGCGCATCTTCACGCCGAGACAGTGTCC	290
Db	591	ATGATCTGGCTACGGGTGCGTGGTGGCGCATCTTCTCGCGGACAGTGTCTCC	532
Qy	291	GGGATCAACCGCTTCGTGGCGCTCTTGGCGGTGGCGCTCTCTCTCTCTCTCTCTCTCTCT	350
Db	531	GGGATCAACCGCTTCGTGGCGCTCTTGGCGGTGGCGCTCTCTCTCTCTCTCTCTCTCTCT	472
Qy	351	ACCAACGACCCCTTCGCATGAACCTGCGCTTCTCTGGCGCGGACACGCTGCGAAGGTG	410
Db	471	ACCAACGACCCCTTACACCATGAACCTGCGTTCATCGCGCGGACACGCTGCGAAGTCT	412
Qy	411	GCCTGCTTCGCGCTGCTGGCGCTTCCGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	470
Db	411	ATGGTCTGGCGATGCT-----CACGCGTGGAGCCACCTCAGCGCGCGGCGAGCC	360
Qy	471	CTCGACTGGAGCATCAGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	530
Db	359	CTGAGTGGACCATCAGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	300
Qy	531	ATCCCGCTCTCGAGGCGATGTACGCGCGCTGCTCGCGGCGCACGCTCATGTTCAGGTTC	590
Db	299	ATCCCCCTCTCAAGGGCATGTACGCGGACTTCT---CCGGCAGGCTCATGTTCAGATC	243
Qy	591	GTCGCTCTCAGTGCATCATCTGCTACAGCTCATGCTCTCTCTCTCTCTCTCTCTCTCTCT	650
Db	242	GTCGCTCTCAGTGCATCATCTGCTACAGCTCATGCTCTCTCTCTCTCTCTCTCTCTCTCT	183



Db 1426 AGTCTAATCGTCTCATATGGGCTCTTGTGCTTACCGGTGGCATGTGGCTATGCCCAA 1485  
Qy 1548 ATCAATGCCCGGTCGATTTGATCCTGTGGAAGCGGGTCTCGGATGCGCATGTTTCAGC 1607  
Db 1486 ATATTACAACAATCCATCTCCATATCTCAGATGCTGGTCTTGGAAATGGCTATGTTTCAGC 1545  
Qy 1608 CTAGGCTGTTTCATGGGCTGCAGCGGAGGATCATCGCGTGGGAAACAAGCTGGCGGCC 1667  
Db 1546 TTAGGTTTATTCATGGCACTTCAACCAAAATCATGCTGTGGGAATCTGTGCGCAAG 1605  
Qy 1668 ATCCGATGGGCTCGGTTGTCGCGAGGCCCGCGTCAAGCGCGCGCTCCATGCGC 1727  
Db 1606 TTTGCCATGGCGGTTCAGATTATTAACCGTCCGGCCATCATGGGTGTGGTGGATGTC 1665  
Qy 1728 GTCGGTCTGCGGGGCTCTCTCCATCGCATGCTCCAGGCTGCTGCGCTCAGGG 1787  
Db 1666 ATTGGCTTACAGCGGAGCTTCTCCGTATAGCAATCGTTAGGCTGGGTGCTCAAGGA 1725  
Qy 1788 ATCGTGCGGTTGCTGTCGCGAGGCCCGCGTCAAGCGCGCTCCATCCGACATCCCTGAGCAGCG 1847  
Db 1726 ATAGTTCGGTTGTGTTGGCAAGAGTACAAATGTCATCCACGATTCCTAAGCACTGG 1785

## RESULT 15

US-10-424-599-45725  
; Sequence 45725, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kowalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 45725  
; LENGTH: 3020  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_141290C.1  
US-10-424-599-45725

Query Match 11.5%; Score 319.8; DB 12; Length 3020;  
Best Local Similarity 53.5%; Pred. No. 1e-67;  
Matches 951; Conservative 0; Mismatches 717; Indels 111; Gaps 9;  
Qy 161 ATCCGCGGCATGATCACCGGCTGGAGCTTACCACTGCTGACGGGGTGGTGGCGCT 220  
Db 745 ACCCAAAATATGATCACCTTAACAGACTTCTACCATGTCATGACTGCAATGGTGCCACT 804  
Qy 221 GTACGTGGCCATGACCTGGCGGTACGGTCCGCTCGGTGGTGGCGCATCTTCAGCGCGGA 280  
Db 805 CTATGTGGCCATGATACCTATAGCTATGGTCACTGAGTGGTGGGAATTTCTCCCTGA 864  
Qy 281 CCAGTCTCCGGGATCAACCGCTTCGTGGCGCTTCGCGCTTCGCGCTGCGCTCTCTCTCCCA 340  
Db 865 CCAATGCTCTGGCATCAACCGGTTTGTGGCACTCTTTTGCAGTGGCTCTCTCTCTCTCCCA 924  
Qy 341 CTTTATCTCCACACGACCCCTTCGCCATGAACCTGCGCTTCTGGCGCGGACACGCT 400  
Db 925 CTTTATAGCTCCAAACACCTTATGATGAATGAACCTGAGTTCTTAGTGTGTGACACCT 984  
Qy 401 GCAGAAGTGGCGGCTCTCGCGCTGCTGGCGCTGCGCTCCCGCGCTCTCTCTCTCTCTCT 460  
Db 985 TCAAAAGATCAATACTAGTCTCTCTTGCAGT-----TTGGAGCAACATCACCA 1035  
Qy 461 CCGCTCGGGCTCGAGTGGAGCATCAGCTTCTCTCCCTCTCCAGCTCCCAACAGCT 520  
Db 1036 AAGGGGTGTGTTGGAAATGGGCCATAACCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 1095

Qy 521 CGTATGGGCATCCCGCTGCTCGAGGCATGTACGGCGCTGCTCGCGCGGACCGCTCAT 580  
Db 1096 GGTATGGGCATCCCGCTGCTCGAGGCATGTATGGTACTCT---CAGGAGCGCTCAT 1152  
Qy 581 GGTCCAGGTGCTGCTGCTCCAGTGTGATCATCTGGTACACGCTCATGCTCTTCTCTTGA 640  
Db 1153 GGTCAAAATGTGGTCTCCAGTGCATCATTTGGTACACCTTGATGCTCTTCTTGTGA 1212  
Qy 641 GTACCGCGCGCGCGCTGCTGCTCGACCAAGTTCGCCGACGCGCGCGCGCTGCTCAT 700  
Db 1213 GTTTAGAGGTGCGAATGCTCATCTCTGAGAGTTCCTGACACTGCTGCTC---TCCAT 1269  
Qy 701 CGTCTCTCTCCGCTGCACTCCGACGCTGCTCTGCTGCTGCTC----- 739  
Db 1270 TGTCTCCATCCATGTGGACTCTGTATGTCTATGTGATGATGATGATGATGATGATGATG 1329  
Qy 740 -----CAGGGGGAGCTCGAGCTCGAGCGCGAGCCGAGCGCGGTGCGCGCGC 787  
Db 1330 TGAAGCTGAGATCAAGGAAGATGGTAAACTCCATGCTCATGCTGAGGAATCCAAATGCTTC 1389  
Qy 788 CGCGCGCTGCTCTCTCCGCGCGCGGAGCGCGGGCGGTGGCGGTGCTGCTGCTGCTGCTGCT 847  
Db 1390 AAGATCAGACATCTTCTCTAGAAAGTCTCAGGCTCTCTCTCCACCATCTCCAGCGCTTC 1449  
Qy 848 CACAGCTCGCTCCGAGCGCGCTGCTGCACTCGCACTCCAGACCATCGAGCGCGC 907  
Db 1450 CAACCTTACCAATGCTGAGATATACTCTTTGCAATC---CTCTAGGAACCTCTACCGCGAG 1506  
Qy 908 TGTGTCCACCTCTCCGCGCTGGAGATCTACTGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCT 962  
Db 1507 AGGCTCTAGTTTCAACCACTGATTTCTATCTCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1566  
Qy 963 ACCCGCGCGGTCCAGCTTCAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1022  
Db 1567 AAATTTTGTGCTCTGATTTTATGGCTTTTACGCTTCAAGAGGCGCAACTCCAAGGCGC 1626  
Qy 1023 AAGGAGCGGAGGAGCGCGCGGAGAGAGAGGCGCATGCGCGCGCGCGCGCGA 1082  
Db 1627 TTCTAATATGATGAAGTGTGGGAAGCCAAAGTTTCATACCATGCTGCTGCTGCTGCTGCTGCT 1686  
Qy 1083 GGACACTCGCGCAGCGCGCGCGCTGCGCGCTGCGCGCTCAAGAGGAAGCACTGCACTG 1142  
Db 1687 TGGGCACTACCTCGACCTAACCTTGGCATGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1746  
Qy 1143 CTCGCTGAGCTCCAGCGCTGCGCGCTGCTGCGAGCGCGCGCGCTGCGACCTCTCTGCGC 1202  
Db 1747 TGTGCTAATGCTAATGCGCAAGAGGCTTAATGGGCGGCTCAGCTGAGAGCTGAGAGTGG 1806  
Qy 1203 GCGCGCGCGCTGACCATGCGCGCTCTCGCGCA--AGAGCGCCAGGCTTACGACGTA 1261  
Db 1807 GAATAGGAGCTTCTATGTTTGGAGTTCAAGTCTTCAAGTCTTCAAGTCTTCAAGTCTTCAAGTCT 1866  
Qy 1262 CGGCGCGAGCACTA-----CAGCAGAGGAGCAAGCAAGCGCGCGCGCGCG 1308  
Db 1867 TGTGCTCCTGATGATGAGAGGCTCATGATCAGAAAGAGTCAAAATGATGATGATGATGATGATG 1926  
Qy 1309 CGGACAAAGCGCGCGCGCTGCTGAGCTGCGAGCTGGGTCCAACTC----- 1351  
Db 1927 AGGAAAGTGGAGATATCATAGGCACTCAAGAGACTACCTAGAGAAAGATGAGTT 1986  
Qy 1352 -----GACGCGCGAGCTGTACCCCAAGGACGACGCGGAGGG 1387  
Db 1987 CAGCTTTGGGAATAGAGAAATGGACAGGAGATGATCAGCTTGAAGGTGAGAGGTTGG 2046  
Qy 1388 GAGGCGCGCGGCTGCGCTGCGCGCGCGAGCGGTGATGAGCGGCTCATCTCTCATCAT 1447  
Db 2047 AGATGGAAACCAAAACCATGCTCCAGAAAGTGTGATGACAGGCTTATATTGATTAT 2106  
Qy 1448 GGTGTGGAGAGCTGATCCGGAACCCCAACCTACTCTCAGCTCATCGGCGTCTG 1507  
Db 2107 GGTGTGGAGAAACTCATCAGAAACCCCAACCTACTCTAGCTTAATTTGGTCTCACTTG 2166



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2004, 06:21:43 / Search time 19 Seconds  
(without alignments)  
1570.326 Million cell updates/sec

Title: US-10-030-884-14

Perfect score: 2910

Sequence: 1 MITALDLYHVLTAVPLVLA.....DILSTAYGPITSHGFITCHS 573

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	139	4.8	302	1 YD82 METHH	O27435 methanobact
2	134	4.6	512	1 FUS BOVIN	Q28009 bos taurus
3	128.5	4.4	308	1 Y311 METJA	Q84337 methanococ
4	126.5	4.3	333	1 HME2 HUMAN	P19622 homo sapien
5	126	4.3	526	1 FUS HUMAN	P35637 homo sapien
6	124.5	4.3	623	1 PNT1 DROME	P51022 drosophila
7	123	4.2	1147	1 MY8B ACACA	P19706 acanthamoeb
8	119.5	4.1	1013	1 PMPG CHLTR	O84879 chlamydia t
9	119	4.1	718	1 MDC6 RHIME	P56949 rhizobium m
10	117	4.0	320	1 FUS MOUSE	P13054 drosophila
11	116	4.0	518	1 KNRL DROME	P46870 chlamydomon
12	115	4.0	647	1 KLP1 CHLRE	Q854h3 bifidobacte
13	114.5	3.9	776	1 FTSK BIFLO	Q8wx85 homo sapien
14	112	3.8	969	1 CCG8 HUMAN	P33244 sorghum bic
15	110.5	3.8	425	1 FTF1 DROME	O54839 mus musculu
16	110	3.8	1043	1 GRP2 SORBI	P13666 cryotolagus
17	109	3.7	168	1 BOND MOUSE	P40146 rattus norv
18	109	3.7	688	1 SRC4 RABIT	Q10707 mycobacteri
19	109	3.7	908	1 CYA8 RAT	P17151 human cytom
20	109	3.7	1248	1 YK98 MYCTU	P23813 mus musculu
21	108.5	3.7	491	1 EP84 HCMVA	P40145 homo sapien
22	108.5	3.7	684	1 HXDB MOUSE	Q92840 chlamydia p
23	108	3.7	323	1 HXDB HUMAN	P13709 drosophila
24	108	3.7	1251	1 Y443 CHLPN	O83983 treponema p
25	107.5	3.7	417	1 FSH DROME	O95029 mesocricetu
26	107	3.7	2038	1 GATA TREPA	O86164 chlamydia p
27	106.5	3.7	506	1 BET3 MESAU	Q09624 caenorhabdi
28	106	3.6	367	1 PM11 CHLPN	O75444 homo sapien
29	105.5	3.6	928	1 YS89 CAEEL	Q9C019 homo sapien
30	105.5	3.6	3178	1 MAF HUMAN	P07875 bacterioph
31	105	3.6	403	1 BHB3 HUMAN	
32	105	3.6	482	1 VGB3 BPT2	
33	104.5	3.6	262	1	

Query Match

4.8%; Score 139; DB 1; Length 302;

## ALIGNMENTS

### RESULT 1

ID	YD82 METHH	STANDARD;	PRT;	302 AA.
AC	O27435;			
DT	15-JUL-1998 (Rel. 36, Created)			
DT	15-JUL-1998 (Rel. 36, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Hypothetical protein MTH1382.			
GN	MTH1382.			
OS	Methanobacterium thermoautotrophicum.			
OC	Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;			
OC	Methanobacteriaceae; Methanothermobacter.			
OX	NCBI_TaxID=187420;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Delta H.			
RX	MEDLINE=98037514; PubMed=9371463;			
RA	Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,			
RA	Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,			
RA	Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,			
RA	Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,			
RA	Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,			
RA	McDougall S., Shiner G., Goyal A., Pietrowski S., Church G.M.,			
RA	Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.			
RT	*Complete genome sequence of Methanobacterium thermoautotrophicum			
RT	deltaH: functional analysis and comparative genomics."			
J. Bacteriol.	179:7135-7155(1997).			
CC	- - SUBCELLULAR LOCATION: Integral membrane protein (Potential).			
CC	- - SIMILARITY: TO M.JANVASCHII MJ1031.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	EMBL; AB000901; AAB85859.1; -			
DR	PIR; H69050; H69050.			
DR	InterPro; IPR004776; Auxin_eff.			
DR	Pfam; PF03547; Auxin_eff; 1.			
DR	TIGRFA; TIGR00946; 2a69; 1.			
KW	Hypothetical protein; Transmembrane; Complete proteome.			
FT	HYPOTHETICAL 10 30 POTENTIAL.			
FT	TRANSMEM 65 85 POTENTIAL.			
FT	TRANSMEM 102 122 POTENTIAL.			
FT	TRANSMEM 130 150 POTENTIAL.			
FT	TRANSMEM 162 182 POTENTIAL.			
FT	TRANSMEM 190 210 POTENTIAL.			
FT	TRANSMEM 224 244 POTENTIAL.			
FT	TRANSMEM 251 271 POTENTIAL.			
FT	TRANSMEM 282 302 POTENTIAL.			
SQ	SEQUENCE 302 AA; 9D3BLBA617AD42C0 CRC64;			

Query Match

4.8%; Score 139; DB 1; Length 302;



	Best Local Similarity	26.0%;	Pred. No.	0.013;	
	Matches	52;	Conservative	40;	Mismatches 80; Indels 28; Gaps 8
QY	379	GADKG---	GPTLSKLGSNSTAQLYPKDDG----	EGRAAAVAMPPASVM--	TRLILIM--- 426
DB	91	GLDSGKTWGLTVA	AAMNMSGFLGYPVTAGIFGSGLVRAIFYDTGTTLMTFTSLGLLSHI	150	
QY	427	-----	VVRKLIRNPNTYSSLI	GVWSLVSYSRWGIEMPAIIARS-	SILSDAGLGMAMEF 478
DB	151	SGGESGRINRKA	RVFFPPPWAFLLVGFNL-----	WG--LPTGIAGITILGYLSGANVPLMI	204
QY	479	SIGLSFWALQPRI	IACGNKLAIAIANGVFVAGPAMAASAIVAGLRVGLLHIAIVQAALPQ	538	
DB	205	SIGL--	TLLFRHSVDATIVSGLRLISPLMAAGISVYLAFRGINSFVTVEASMP	262	
QY	539	GIVPFVFAKEYGV	HDPDILTST	558	
DB	263	AMLAAVLATEND	LVDVLVSS	282	
 RESULT 2 FUS_BOVIN STANDARD; PRT; 512 AA. 					
ID	FUS_BOVIN	STANDARD;	PRT;	512	AA.
AC	Q28009;				
DT	01-NOV-1997	(Rel. 35, Created)			
DT	01-NOV-1997	(Rel. 35, Last sequence update)			
DE	15-MAR-2004	(Rel. 43, Last annotation update)			
DE	RNA-binding protein FUS	(Pigpen protein).			
GN	FUS.				
OS	Bos taurus (Bovine).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;				
OX	Bovidae; Bovinae; Bos.				
OX	NCBI_TaxID=9913;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Aorta;				
RX	MEDLINE=96175600; PubMed=8631501;				
RT	Alliegro M.C., Alliegro M.A.;				
RA	"A nuclear protein regulated during the transition from active to				
RT	quiescent phenotype in cultured endothelial cells.";				
RL	Dev. Biol. 174:288-297(1996).				
RN	[2]				
RP	CARBOHYDRATE-BINDING DOMAIN.				
RX	MEDLINE=20160719; PubMed=10694442;				
RA	Alliegro M.C.;				
RT	"A C-terminal carbohydrate-binding domain in the endothelial cell				
RT	regulatory protein, pigpen: new function for an EWS family member.";				
Exp. Cell Res.	255:270-277(2000).				
-!	FUNCTION: Binds both single-stranded and double-stranded DNA and				
CC	promotes ATP-independent annealing of complementary single-				
CC	stranded DNAs and D-loop formation in superhelical double-stranded				
CC	DNA. May play a role in maintenance of genomic integrity (By				
CC	similarity).				
CC	-! SUBUNIT: Component of nuclear riboprotein complexes. Binds SF1 (By				
CC	similarity).				
CC	-! SUBCELLULAR LOCATION: Nuclear, exhibits diffuse staining				
CC	throughout (excluding nucleoli), together with a small number of				
CC	intensely stained focal points, or granules, and punctate staining				
CC	along the nuclear envelope.				
CC	-! DOMAIN: The C-terminal domain binds carbohydrates.				
CC	-! SIMILARITY: Contains 1 RANBP2-type zinc finger.				
CC	-! SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.				
CC	-! SIMILARITY: BELONGS TO THE TET FAMILY OF RNP PROTEINS.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation				
CC	the European Bioinformatics Institute. There are no restrictions on it				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce				
CC	or send an email to license@isb-sib.ch).				
CC	-----				

DR	EMBL; U26024; AACI3543.1; --
DR	HSSP; P09651; IHAL
DR	InterPro; IPR000504; RNA_rec_mot.
DR	InterPro; IPR001876; Znf_RanGDP.
DR	Pfam; PF00076; xrm; 1.
DR	Pfam; PF00641; zf-RanBP; 1.
DR	SMART; SM00360; RRM; 1.
DR	SMART; SM00547; ZNF RBZ; 1.
DR	PROSITE; PS50102; RRM; 1.
DR	PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
DR	PROSITE; PS01358; ZF_RANBP2_1; 1.
DR	PROSITE; PS50199; ZF_RANBP2_2; 1.
KW	RNA-binding; DNA-binding; Nuclear protein; Repeat; Zinc-finger; Zinc;
KW	Metal-binding.
FT	DOMAIN 1 164 GLN/GLY/SER/TYR-RICH.
FT	DOMAIN 165 253 GLY-RICH.
FT	DOMAIN 271 357 RNA-BINDING (RRM).
FT	DOMAIN 357 512 ARG/GLY-RICH.
FT	ZN_FING 408 439 RANBP2-TYPE.
SQ	SEQUENCE 512 AA; 25240 MW; 3652329C04F1386 CRC64;

Query Match 4.6%; Score 134; DB 1; Length 512;  
Best Local Similarity 23.6%; Pred.No.0.056;  
Matches 56; Conservative 27; Mismatches 78; Indels 76; Gaps 11;

QY	197 EAEIPDGVAGACAVSSRGDGAGRVRVTYRKSTSRSEAAASHSHSQTMQPRVSNLSG----	252
DB	69 QSAPOQVGSAAGYGSS-----QSSQSSYGGQQSYVGYGQ--QPAPSSTSGSYGS	115
QY	253 -----VVIYLOSSRNPTPGSSFNHADFFENVGAALKGGG	288
DB	116 SSOSSGYGQGQGGYGGQSGYGGQQSQQQSYNPP-QQS-:-YQQSQYNSSGGCGGGGG	172
QY	289 GAAG-DREKACGGGGGGHGFQPAVAVPARKDLHLVLWSSASPVSERAHVHFGAGG	347
DB	173 GSYGQDQPSMSSGGGGGYGNQDQSGYGGGQQD-----RGRGRGGGGG	217
QY	348 ADHADVLAKQAQAYDEYGRDDYSRTKNNGSGGADKGPTSLKLGNSSTAQLYPKDDG	404
DB	218 YN-----RSSGGYEPRGGG-GGGGRRGGMGSDRGG--FNKFGG-----FRDQG	258

RESULT 3

YA31_METJA	
ID_YA31_METJA	STANDARD; PRT; 308 AA.
AC_Q58437;	
DT_15-JUL-1998	(Rel. 36, Created)
DT_15-JUL-1998	(Rel. 36, last sequence update)
DT_16-OCT-2001	(Rel. 40, Last annotation update)
DE_Hypothetical protein MJ1031.	
DN_MJ1031.	
OS_Methanococcus jannaschii.	
OC_Archaea; Euryarchaeota; Methanococci; Methanococcales;	
OC_Methanocaldococcaceae; Methanocaldococcus.	
OX_NCBI_TaxID=2190;	
[1]	
RN_SEQUENCE FROM N.A.	
RC_STRAIN=JAL-1 / DSM 2661 / ATCC 43067;	
RC_MEDLINE=96337999; PubMed=8688087;	
RA_Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,	
RA_Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,	
RA_Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,	
RA_Ovcrbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,	
RA_Scott J.L., Geoghagen S.N.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,	
RA_Utterback T.R., Kelley J.M., Peterson J.D., Sadov P.W., Hanna M.C.,	
RA_Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,	
RA_Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;	
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus	
RA_jannaschii";	
RL_Science 273:1058-1073(1996).	
RC_-!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).	
-!-, SIMILARITY: TO M.THERMOAUTOTROPHICUM MTH1382.	
CC-----	

[illegible]

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))

EMBL; U67545; AAB90035.1; --  
 PIR; F64428; F64428.  
 TIGR; M1031; --  
 InterPro; IPR004776; Auxin\_eff.  
 Pfam; PF03547; Auxin\_eff; 1.  
 TIGRPFAMs; TIGR00946; 2a69; 1.  
 KW Hypochemical protein; Transmembrane; Complete proteome.  
 FT TRANSMEM 5 26 POTENTIAL.  
 FT TRANSMEM 31 51 POTENTIAL.  
 FT TRANSMEM 63 83 POTENTIAL.  
 FT TRANSMEM 100 120 POTENTIAL.  
 FT TRANSMEM 128 148 POTENTIAL.  
 FT TRANSMEM 162 182 POTENTIAL.  
 FT TRANSMEM 195 215 POTENTIAL.  
 FT TRANSMEM 221 241 POTENTIAL.  
 FT TRANSMEM 257 277 POTENTIAL.  
 FT TRANSMEM 287 307 POTENTIAL.  
 SQ SEQUENCE 308 AA; 33770 MW; 02A915806E4441AE CRC64;

Query Match 4.4%; Score 128.5; DB 1; Length 308;  
 Best Local Similarity 22.5%; Pred. No. 0.072;  
 Matches 46; Conservative 46; Mismatches 87; Indels 25; Gaps 5;

QY 383 GGTLSKLSNSTAQLYPKDDG---EGRAAVAMPASVTRLLINW-----R 429  
 DB 97 GGLVLSMGNLTFGLYFVALGMEGLAFAICDLGGVFATMLLGYFVGIRFGKGRDK 156  
 QY 430 KLIRNPTYSLLGVVMSLVSRNGIEM---PAIARISILSDAGLGMAMFSLGLFWAL 485  
 DB 157 SILKMAKPPPLTGILSIILVFFGKLYIPSFILKSLNLSATVPLNWSLGL--SL 214  
 QY 487 QRIACGNKLAIAAGVRFVAGPAVMAAASIAVGLRGLVLIHIAVQAALPQGVIPVFA 546  
 DB 215 SPKALFGVFWGIIASIRFIVSPATATLSELINIKLEKNVLLVSSMPSAMMTLVLG 274  
 QY 547 KEYGVHPDILSTAYGPITSHGPI 570  
 DB 275 TLVELDIKLIASSI-----FIT 291

## RESULT 4

HME2\_HUMAN STANDARD; PRT; 333 AA.  
 AC P19622;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Homeobox protein engrailed-2 (Hu-En-2).  
 GN EN2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93185339; PubMed=1363401;  
 RA Logan C., Hanks M.C., Noble-Topham S., Nallainathan D.,  
 RA Provart N.J., Joyner A.L.;  
 RT "Cloning and sequence comparison of the mouse, human, and chicken  
 RT engrailed genes reveal potential functional domains and regulatory  
 RT regions."  
 RL Dev. Genet. 13:345-358(1992).  
 RN [2]  
 RP REVISION TO 229.  
 RA Logan C., Hanks M.C., Noble-Topham S., Nallainathan D., Provart N.J.,

Joyner A.L.;  
 Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 [3]

RA SEQUENCE OF 230-333 FROM N.A.  
 RN MEDLINE=89233109; PubMed=2565873;  
 RP Poole S.J., Law M.L., Kao F.T., Lau Y.-F.C.;  
 RA "Isolation and chromosomal localization of the human En-2 gene."  
 RT Genomics 4:225-231(1989).  
 CC -|- SUBCELLULAR LOCATION: Nuclear.  
 CC -|- SIMILARITY: Belongs to the engrailed homeobox family.  
 CC -|- SIMILARITY: Contains 1 homeobox domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))

EMBL; L12701; AAA53504.2; --  
 EMBL; L12700; AAA53504.2; JOINED.  
 EMBL; J03066; AAF68670.1; --  
 FIR; E48423; E48423.  
 HSP; P02836; 3HDD.  
 DR TRANSFAC; T02019; --  
 DR MIM; 131310; --  
 DR GO; GO:0007275; P:development; TAS.  
 DR InterPro; IPR000747; Engrailed.  
 DR InterPro; IPR001356; Homeobox.  
 DR InterPro; IPR000047; HTH\_lambrepreser.  
 DR Pfam; PF00046; homeobox; 1.  
 DR PRINTS; PR00026; ENGRAILED.  
 DR PRINTS; PR00024; HOMEBOX.  
 DR PRINTS; PR00031; HTHREPRESSR.  
 DR ProDom; PD000010; Homeobox; 1.  
 DR SMART; SM00389; HOX; 1.  
 DR PROSITE; PS00027; HOMEBOX\_1; 1.  
 DR PROSITE; PS00071; HOMEBOX\_2; 1.  
 DR PROSITE; PS00033; ENGRAILED; 1.  
 KW Homeobox; DNA-binding; Developmental protein; Nuclear protein.  
 FT DNA\_BIND 244 303 HOMEBOX.  
 SQ SEQUENCE 333 AA; 34210 MW; ACF5399B83D6257 CRC64;

Query Match 4.3%; Score 126.5; DB 1; Length 333;  
 Best Local Similarity 28.1%; Pred. No. 0.11; Mismatches 18; Indels 57; Gaps 10;

Matches 62; Conservative 18; Mismatches 84; Indels 57; Gaps 10;  
 QY 167 DQFPDGAASIVSRVSDVVSVLARGDVELEAEAPDGAVAGAVSSRG-GDAGRVRVTVRK 225  
 DB 5 DPKPGEAAAV-----EGORQPESSPGGGGGGGSSPGADTGRRALMLP 51  
 QY 226 STSSRSEACSHSHSQTMQPRVNSLGVETYSLSQSRNPTPRGSSFNHADFFNIVGAAK 285  
 DB 52 AV---LQAPGNHQPH-----RITNFFDNLIRPEFGR-----RKDAGTCCAGAGG 94  
 QY 286 GGGGAGDE--EKACGGGGGGHSPQPAVAVPAKRXDLHMLVWSSASVPSRAAVHVF 343  
 DB 95 GRGGGAGGEGGASGAEGGGAGGSEQ-----LLSGSGSRPRQNPFCAP-- 137  
 QY 344 GAGGADHADVLAKQAQAYDEVGRDDYSSRTKNGSGGADKGG 384  
 DB 138 GAGG-----PLPAAGS---DSPGDEGGGSKTSLHGNAKGG 171

## RESULT 5

FUS\_HUMAN STANDARD; PRT; 526 AA.  
 ID FUS\_HUMAN  
 AC P3537;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE RNA-binding protein FUS (Oncogene FUS) (Oncogene TLS) (Translocated in  
DE liposarcoma protein) (POMP75) (75 kDa DNA-pairing protein).  
GN FUS OR TLS.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Sutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RN SEQUENCE FROM N.A. (ISOFORM LONG).  
RX MEDLINE=93288139; PubMed=8510758;  
RA Crozat A., Aman P., Mandahl N., Ron D.;  
RT "Fusion of CHOP to a novel RNA-binding protein in human myxoid  
RT liposarcoma";  
RL Nature 363:640-644(1993).  
RN [2]  
RN SEQUENCE FROM N.A. (ISOFORM SHORT).  
RX MEDLINE=93350637; PubMed=7503811;  
RA Rabbitts T.H., Forster A., Larson R., Nathan P.;  
RT "Fusion of the dominant negative transcription regulator CHOP with a  
RT novel gene FUS by translocation t(12;16) in malignant liposarcoma";  
RL Nat. Genet. 4:175-180(1993).  
RN [3]  
RN SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT).  
RX MEDLINE=99013873; PubMed=9795213;  
RA Morohoshi F., Ootsuka Y., Azai K., Ichikawa H., Mitani S.;  
RA Munakata N., Ohki M.;  
RT "Genomic structure of the human RBP56/HTAF1168 and FUS/TLS genes";  
RL Gene 221:191-198(1998).  
RN [4]  
RN SEQUENCE FROM N.A. (ISOFORM LONG).  
RP TISSUE=Lung, and Lymph;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zerbahn B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Diatchenko M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loughlano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Munz D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalios D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [5]  
RN SEQUENCE OF 235-244; 307-312; 335-345 AND 349-357, AND  
RN CHARACTERIZATION.  
RX MEDLINE=20036580; PubMed=10567410;  
RA Baechtold H., Kuroda M., Sok J., Ron D., Lopez B.S., Akhmedov A.T.;  
RT "Human 75-kDa DNA-pairing protein is identical to the pro-oncoprotein  
RT TLS/FUS and is able to promote D-loop formation";  
RL J. Biol. Chem. 274:34337-34342(1999).  
RN [6]  
RN SEQUENCE OF 265-276; 317-331 AND 335-357, AND IDENTIFICATION.  
RX MEDLINE=99369251; PubMed=10442642;  
RA Bertrand P., Akhmedov A.T., Delacote F., Durrbach A., Lopez B.S.;  
RT "Human POMP75 is identified as the pro-oncogene TLF/FUS: both  
RT POMP75 and POMP100 DNA homologous pairing activities are associated  
RT to cell proliferation";  
RL Oncogene 18:4515-4521(1999).  
RN [7]  
RN CHROMOSOMAL TRANSLOCATION.  
RX MEDLINE=94243799; PubMed=8187069;  
RA Ichikawa H., Shimizu K., Hayashi Y., Ohki M.;  
RT "An RNA-binding protein gene, TLS/FUS, is fused to ERG in human

RT myeloid leukemia with t(16;21) chromosomal translocation.";  
RL Cancer Res. 54:2865-2868(1994).  
RN [8]  
RN INTERACTION WITH SP1.  
RX MEDLINE=98325009; PubMed=9660765;  
RA Zhang D., Paley A.J., Childs G.;  
RT "The transcriptional repressor ZFM1 interacts with and modulates the  
RT ability of EMS to activate transcription";  
RL J. Biol. Chem. 273:18086-18091(1998).  
CC -1- FUNCTION: Binds both single-stranded and double-stranded DNA and  
CC promotes ATP-independent annealing of complementary single-  
CC stranded DNAs and D-loop formation in superhelical double-stranded  
CC DNA. May play a role in maintenance of genomic integrity.  
CC -1- SUBUNIT: Component of nuclear riboprotein complexes. Interacts  
CC with ILF3 and SF1.  
CC -1- SUBCELLULAR LOCATION: Nuclear.  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=Long;  
CC IsoId=P35637-1; Sequence=Displayed;  
CC Name=Short;  
CC IsoId=P35637-2; Sequence=VSP\_005798;  
CC -1- TISSUE SPECIFICITY: Ubiquitous.  
CC -1- DISEASE: Involved in a form of malignant myxoid liposarcoma  
CC through a chromosomal translocation t(12;16) (q13;p11) that  
CC involves D01T3 and FUS.  
CC -1- DISEASE: Involved in a form of acute myeloid leukemia (AML)  
CC through a chromosomal translocation t(16;21) (p11;q22) that  
CC involves FUS and ERG.  
CC -1- SIMILARITY: Contains 1 RANBP2-type zinc finger.  
CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.  
CC -1- SIMILARITY: BELONGS TO THE TET FAMILY OF RNP PROTEINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC EMBL: S62140; AB27102.1; -;  
CC EMBL: S62138; AB27103.1; ALT\_SEQ.  
CC EMBL: X71427; CAAS0558.1; ALT\_SEQ.  
CC EMBL: X71428; CAAS0559.1; ALT\_SEQ.  
CC EMBL: AF071213; AAC35285.1; -;  
CC EMBL: AF071213; AAC35284.1; -;  
CC EMBL: BC000402; AAH00402.1; -;  
CC EMBL: BC002459; AAH02459.1; -;  
CC PIR: S33799; S33799.  
CC HSRP; P09651; IHA1.  
CC Genew; HGNC:4010; FUS.  
CC GK; P35637; -;  
CC MIM; 137070; -;  
CC MIM; 151900; -;  
CC GO; GO:0005634; C:nucleus; TAS.  
CC GO; GO:0003723; F:RNA binding; TAS.  
CC InterPro; IPR000504; RNA\_rec\_mot.  
CC InterPro; IPR001876; Znf\_RanGDP.  
CC Pfam; PF00076; rrm; 1.  
CC Pfam; PF00641; zf-RanBP; 1.  
CC SMART; SM00360; RRM; 1.  
CC SMART; SM00547; Znf\_RBZ; 1.  
CC PROSITE; PS50102; RRM; 1.  
CC PROSITE; PS00030; RRM\_RNP\_1; FALSE\_NEG.  
CC PROSITE; PS01358; ZF\_RANBP2\_1; 1.  
CC PROSITE; PS50199; ZF\_RANBP2\_2; 1.  
CC Proto-oncogene; RNA-binding; DNA-binding; Nuclear protein; Repeat;  
CC Alternative splicing; Chromosomal translocation; Zinc-finger; Zinc;  
CC Metal-binding.  
KW DOMAIN 1 165 GLN/GLY/SER/TYR-RICH.  
FT DOMAIN 166 267 GLY-RICH.  
FT DOMAIN 285 371 RNA-BINDING (RRM).

FT DOMAIN 371 526 ARG/GLY-RICH.  
 FT ZN FING 422 453 RANP2-TYPE  
 FT SITE 266 BREAKPOINT FOR TRANSLOCATION TO FORM  
 FT VARSPLIC 64 65 FUS/TLS-CHOP ONCOGENE.  
 FT CONFLICT 338 338 TG -> S (in isoform Short).  
 FT SEQUENCE 526 AA; 53426 MW; 88C8E263B7905549 CRC64;  
 Query March 4.3%; Score 126; DB 1; Length 526;  
 Best Local Similarity 27.1%; Pred. No. 0.21;  
 Matches 56; Conservative 18; Mismatches 75; Indels 58; Gaps 10;  
 Qy 208 AVSRGGDAGRVTVRKSTSSREACSHSOTMOPRVNLSGVBIYSLQSRNPTPR 267  
 Db 105 APSSTSGSYG-----SSQSASVGPQSGSYSQPSYGGQ--QSYGQQSYNP-PQ 153  
 Qy 268 GSSFNHDFFNIVGAAGKGGGAAGDEKAGCGGGGGHSPQPAVAVPAKRDLEMLVW 327  
 Db 154 G--YQQNQYN-----SSSGGGG-----GGGGGGNYGDDQ----- 181  
 Qy 328 SSSASPSVSEAAVHFGAGGADHADVLAKGAQAYDEYRDDYSRTKNGSGGADKGGPTL 387  
 Db 182 SSMSSG-----GGSGGYGNQDSGGGGGGYGGQDRGGRGGSGGGGGGGGG- 230  
 Qy 388 SKLGSNSTAQLY-PKDDGEGRAAAVAM 413  
 Db 231 --GVNRRSGGYEPGRGGRGGRGGRGGM 254  
 RESULT 6  
 ID PNT1 DROME STANDARD; PRT; 623 AA.  
 AC P51022; P19420;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE ETS-like protein pointed, isoform P1 (D-ETS-2).  
 GN PNT OR ETS8AB OR ETS2 OR CGI7077.  
 OS Drosophila melanogaster (fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 CX NCBI\_TaxID=7227;  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=20196006; PubMed=10731132;  
 RX Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson X.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Boutin J., Brockstein P., Brottier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos J.M., Delcher A., Deng X., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaeser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattet B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster."  
 RL Science 287:2185-2195(2000).  
 RN [3]  
 RP REVISIONS, AND ALTERNATIVE SPLICING.  
 RX MEDLINE=22426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
 RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
 systematic review."  
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
 RN [4]  
 RP SEQUENCE OF 456-613 FROM N.A.  
 RX MEDLINE=92249640; PubMed=1577186;  
 RA Chen T., Bunting M., Karim F.D., Thummel C.S.;  
 RT "Isolation and characterization of five Drosophila genes that encode  
 an ets-related DNA binding domain."  
 RL Dev. Biol. 151:176-191(1992).  
 RN [5]  
 RP SEQUENCE OF 445-603 FROM N.A., TISSUE SPECIFICITY, AND DEVELOPMENTAL  
 RP STAGE.  
 RX MEDLINE=88196618; PubMed=2834248;  
 RA Pribyl L.J., Watson D.K., McWilliams M.J., Ascione R., Papas T.S.;  
 RT "The Drosophila ets-2 gene: molecular structure, chromosomal  
 localization, and developmental expression."  
 RL Dev. Biol. 127:45-53(1988).  
 CC -I- FUNCTION: Required for glial-neuronal cell interactions at the  
 CC ventral midline which are necessary for the proper elaboration of  
 CC commissures in the embryonic CNS.  
 CC -I- SUBCELLULAR LOCATION: Nuclear (Potential).  
 CC -I- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=P1; Synonyms=C;  
 CC IsoId=P51022-1; Sequences=Displayed;  
 CC Name=P2; Synonyms=B;  
 CC IsoId=P51023-1; Sequences=External;  
 CC -I- TISSUE SPECIFICITY: Expressed in a complex dynamic pattern in  
 CC early embryos, including the midline and midline glial cells.  
 CC -I- DEVELOPMENTAL STAGE: Expressed throughout development with lower  
 CC levels during larval development.  
 CC -I- SIMILARITY: Belongs to the ETS family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X69166; CAA48916.1; -

```

DR EMBL; AE003742; AAN13943.1; -.
DR EMBL; M88472; AAC34200.1; -.
DR EMBL; M20408; AAA28521.1; -.
DR PIR; S3167; S3167.
DR PIR; S3168; S3168.
DR HSP; P14921; 2STT.
DR FLYBase; FBgn0003118; pnt.
DR GO; GO:0005026; F-type II transforming growth factor-beta rec. . .; NAS.
DR GO; GO:0006916; P:anti-apoptosis; IGI.
DR GO; GO:0007173; P:EGF receptor signaling pathway; NAS.
DR GO; GO:0007362; P:terminal region determination; IGI.
DR GO; GO:0007179; P:FGFRbeta receptor signaling pathway; NAS.
DR GO; GO:0008293; P:torso signaling pathway; IGI.
DR GO; GO:0007424; P:tracheal system development (sensu Insecta); NAS.
DR InterPro; IPR000418; Ets.
DR InterPro; IPR002341; HSF_ETS.
DR Pfam; PF00178; Ets; 1.
DR PRINTS; PR00454; ETSDOMAIN.
DR SMART; SM00413; ETS; 1.
DR PROSITE; PS00345; ETS_DOMAIN_1; 1.
DR PROSITE; PS00346; ETS_DOMAIN_2; 1.
DR PROSITE; PS00661; ETS_DOMAIN_3; 1.
DR DNA-binding; Nuclear Protein; Developmental protein;
KW Alternative splicing.
FT DOMAIN 38 46 POLY-SER.
FT DOMAIN 47 50 POLY-ASN.
FT DOMAIN 75 88 POLY-HIS.
FT DOMAIN 103 111 POLY-GLN.
FT DOMAIN 119 124 POLY-GLN.
FT DOMAIN 128 134 POLY-GLN.
FT DOMAIN 156 159 POLY-THR.
FT DOMAIN 160 174 POLY-SER.
FT DOMAIN 246 252 POLY-GLN.
FT DOMAIN 264 267 POLY-ASN.
FT DOMAIN 294 298 POLY-ASN.
FT DOMAIN 310 320 POLY-ASN.
FT DOMAIN 323 326 POLY-GLY.
FT DOMAIN 386 389 POLY-ALA.
FT DNA BIND 515 595 ETS-DOMAIN.
FT CONFLICT 467 467 Q -> R (IN REF. 5).
SQ SEQUENCE 623 AA; 66866 MW; P47083D960EB6480 CRC64;

Query Match 4.3%; Score 124.5; DB 1; Length 623;
Best Local Similarity 20.6%; Pred. No. 0.32;
Matches 80; Conservative 55; Mismatches 134; Indels 119; Gaps 15;

QY 161 ARALVLPQFPDGA--AAISIVSFVDSVSVSLARG-----DVELEAEPDGV 203
Db 196 AASLGLGVENDMAFFVGDANAYTDSVNFSSGYNTSNTDRINNSTPPQQCQSQPTV 255
QY 204 AGAGAVSRGSDAGRVTVRKTSRSEREAACHSHSQTWQPRVSNLSGVETYSLOSRN 263
Db 256 NGSGSSSNNNNS--MLPPAVQCGNNNNNTSSNTNNSNN---NNNSGGSNNNSNAGSNN 311
QY 264 PTPRGGSFNFHADFNIVGAAA-----KGGGGAAGDEEXGACGGGGGSHSPQPAVAVPA 317
Db 312 -----NNNNNNINFWAAAIQHLKKEPGT---QNGNIGYGGGSNSQNDPTDLLS 361
QY 318 KRKDLMLVW--SSASPVSEARAHHVFGAGGAD---HADVLAKGAQAYDEYGRDDYSR 372
Db 362 YGLPAHLAAAYGGGSGSGTGGRSS---GGGGSDSYHSTISAQDHQSQQSSG----- 410
QY 373 TNGSGGADKGGPTLSKLGNSNSTAQLPKDDGSEGR----- 407
Db 411 -GNSSGSGSGSGTSGNSGYLDSSEFVGSYAGNRHFHDGYPEFTPYDAQSFQSGMPQPT 469
QY 408 -----AAAVAMPASVMTLLI-----IMWRKLI-----RN 434
Db 470 ANDQWGAHQAHRPAATMSTGLGDKLGLLGGYTTQGGVPCFTGSGPIQLMQFLLELLDKT 529
QY 435 PNTYSSLIGVWVSL-----VSVRWGI 455
Db 530 CQSPISWTGCGWEFKLTPDEVARRWGI 557

```

```

RESULT 7
MYSB ACACA STANDARD; PRT; 1147 AA.
ID MYSB ACACA STANDARD; PRT; 1147 AA.
AC P19706;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Myosin heavy chain IB (Myosin heavy chain IL).
GN MIB OR MIL.
OS Acanthamoeba castellanii (Amoeba).
OC Eukaryota; Acanthamoebidae; Acanthamoeba.
RX NCBI_TaxID=5755;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90080816; PubMed=2511079;
RA Jung G., Schmidt C.J., Hammer J.A. III;
RT "Myosin I heavy-chain genes of Acanthamoeba castellanii: cloning of a
RT second gene and evidence for the existence of a third isoform.";
RL Gene 82:269-280(1989).
CC -!- FUNCTION: Myosin is a protein that binds to F-actin and has ATPase
CC activity that is activated by F-actin.
CC -!- SUBUNIT: Myosin I heavy chain is single-headed. Dimer of a heavy
CC and a light chain. Inability to self-assemble into filaments.
CC -!- DOMAIN: TH.1 binds directly to anionic phospholipid membranes;
CC myosins I could therefore move actin relative to membranes and
CC vice versa. TH.2 and SH3 bind tightly to F-actin; this together
CC with the nucleotide-sensitive site in the head, allows single
CC molecules of myosin I to cross-link actin filaments.
CC -!- MISCELLANEOUS: This organism expresses at least three isoforms of
CC myosin I heavy-chain, encoded by genes MIA, MIB, and MIC.
CC -!- SIMILARITY: Contains 1 myosin-like globular head domain.
CC -!- SIMILARITY: Contains 1 SH3 domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M30780; AAA27708.1; -.
CC PIR; JQ0095; MWAXIB.
CC HSP; P08799; IMND.
CC InterPro; IPR000048; IQ region.
CC InterPro; IPR001609; myosin_head.
CC InterPro; IPR000108; Neu_cyt_fact_2.
CC InterPro; IPR001452; SH3.
CC Pfam; PF00612; IQ; 1.
CC Pfam; PF00663; myosin_head; 1.
CC Pfam; PF00018; SH3; 1.
CC PRINTS; PR00193; MYOSINHEAVY.
CC PRINTS; PR00499; P67PROX.
CC PRINTS; PR00452; SH3DOMAIN.
CC ProDom; PD000355; myosin_head; 1.
CC ProDom; PD000066; SH3; 1.
CC SMART; SM00242; MYSC; 1.
CC SMART; SM00326; SH3; 1.
CC PROSITE; PS00002; SH3; 1.
CC Myosin; ATP-binding; Phosphorylation; Multigene family; SH3 domain.
FT DOMAIN 1 670 MYOSIN HEAD-LIKE.
FT DOMAIN 671 907 TAIL HOMOLOG REGION 1 (TH.1).
FT DOMAIN 908 1089 GLY/PRO/ALA-RICH (TH.2).
FT DOMAIN 1090 1147 SH3.
FT NP BIND 103 110 ATP (POTENTIAL).
SQ SEQUENCE 1147 AA; 124958 MW; B76DE9C076381054 CRC64;

```

Query Match 4.2%; Score 123; DB 1; Length 1147;  
 Best Local Similarity 24.6%; Pred. No. 0.85;  
 Matches 60; Conservative 22; Mismatches 100; Indels 62; Gaps 9;

QY 221 VTVKTSRSEACSHSHSQWQPRVSN-----LSGVEIYLSOSRNPTRGSSF 271  
 Db 861 ITYKAGNQRRLTFKSNESASQAQPSIKKSRANIQIGIATGLPKETDSSPNWTFSGGG 920  
 QY 272 NHADFNIVGAAKGGG-GAAGDEEKGACGGGGHSPQQAQVAVPAKRDHLMLVSSS 330  
 Db 921 GYGG-----GRGGGGGRGAGGRCGFGGGGGGYS-QPVAQAQVAVQPPVAAPSA 974  
 QY 331 -----ASPVSRAAVHVFGAGGADHADVLAQGAQYDEY 364  
 Db 975 GRGGPGMGFGAGRGPGMGFGGPGMGFGGPGMGFGGPGMGFGGPGMGFGGPGMGFG 1028  
 QY 365 GRD-----DYSSRTKNGSGGADGGPTLSK-----LGSNSTAQLYPKDGGEGRAAVAM 413  
 Db 1029 GRGGPGMGFGAGRGPGMGFGGPGMGFGGPGMGFGGPGMGFGGPGMGFGGPGMG 1085  
 QY 414 PPAS 417  
 Db 1086 APAA 1089

RESULT 8  
 PMPG\_CHLTR STANDARD; PRT; 1013 AA.  
 AC 084879;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Probable outer membrane protein pmpG precursor (Polymorphic membrane protein G)  
 DE PMPG OR CT871.  
 GN Chlamydia trachomatis.  
 OS Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCBI\_TaxID=813;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=D/UW-3/Cx;  
 RX MEDLINE=99000809; PubMed=9784136;  
 RA Stephens R.S., Kaiman S., Lemmel C.J., Fan J., Marathe R., Aravind L., Mitchell W.P., Olinger L., Tatusov R.D., Zhao Q., Koonin E.V., Davis R.W.;  
 RA "Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis.";  
 RT Science 282:754-759(1998).  
 RL -1- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies) (Potential).  
 CC -1- SIMILARITY: Belongs to the PMP outer membrane protein family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; AB001360; AAC68469.1; .  
 DR F1R; G71460; G71460.  
 DR PHC1-2DPAGE; 084879; .  
 DR InterPro; IPR006315; Autotransport.  
 DR InterPro; IPR005546; Autotransporter.  
 DR InterPro; IPR003368; Chlamydia\_PMP.  
 DR InterPro; IPR000437; Prok\_lipoprot\_S.  
 DR Pfam; PF03797; Autotransporter; 1.  
 DR Pfam; PF02415; Chlamydia\_PMP; 1.  
 DR TIGRFAMs; TIGR01414; autotrans\_bar1; 1.  
 DR TIGRFAMs; TIGR01376; POMP repeat; 6.  
 DR Outer membrane; Signal; Multigene family; Complete proteome.  
 KW SIGNAL 1 27 POTENTIAL  
 FT CHAIN 28 1013 PROBABLE OUTER MEMBRANE PROTEIN PMPG.  
 SQ SEQUENCE 1013 AA; 107366 MW; F0927743C0A651DD CRC64;

Query Match 4.1%; Score 119.5; DB 1; Length 1013;

Best Local Similarity 21.4%; Pred. No. 1.3;  
 Matches 124; Conservative 65; Mismatches 199; Indels 191; Gaps 30;

QY 81 AVLALL--ALASRGLSSPRALGLDWSITLFSLSLTLPNTLVMG-----IFLLRGMVGAS 131  
 Db 136 SILAVLPAATWNGSQTPPTTSTPSNGTIYKTDL-----LLNNEKFFSYSLVSGDGGAI 192  
 QY 132 SAGTLWQVVVLQCIITWTLMLFLFEYRAARALVLDQPPDGAASIV-SRPVSDSVSLA 190  
 Db 193 DAKSLTVQGISKLCV-----PQNTAQA-----DGGACQVVTFSAMANEAPIA 236  
 QY 191 RGDVELEAPDQVAGAGVSSRGDAGRVVTRKSTSSRSE-AACSHSHSQWQPRVSN 249  
 Db 237 -----FIANVAGVREGGIAAVQDGGQ-----VSSSTSTEDPVVSFRNIAVEFDGNVAR 286  
 QY 250 LSGVEIYLSQ-----SSRNPTRGSSFNHADFFNIVGAA--AKG 286  
 Db 287 VGG-GIYSYGNVAFNLNGKTLFLANNVSPVYIAAEQPT-NQASNTSDNYGDDGGAIFCKN 344  
 QY 287 GGGAAAGDEBK-----ACGGGGGGHSPQQAQVAVPAKRDHLMLVSSS 330  
 Db 345 GAAAGSNNSSVSFDGEGVVFSSNVAAGKG-----AIYAKKLSV-----AN 388  
 QY 331 ASPV-----SRAAAVHVFGAG-----GAHADVLAKG-----AOAYDEYGRDDVSSR 372  
 Db 389 CGPVQFLGNIANDGGAIYLGESGLSLADYGDIIFDGNLKRKTAKENAADVNGVTSSQA 448  
 QY 373 TKNGSGG-----ADKGG-----PTLSKLGNSTAQLYPK--DDGEGRAAAAMVPPAS 417  
 Db 449 ISMGSGGKITTIRAKAGHQIIFNDPIEMANGNNOQASSEPLKINDGEGTGDIVFANG- 507  
 QY 418 VMTRLILIMWRKLRNPNTYSSLGIVVSVYRWGIEMPAIAR-----SISILSDAG 472  
 Db 508 -----NSTLYQNV-----IEQRIVLREKAKLSVNSLSQTG 539  
 QY 473 --LGVAMPSELGLFMALOPRIIACGNKLAIAIANGVRFVAGPAVMAAASIAVGLRGVLLHIA 530  
 Db 540 GSLYWEAGSTLDFVTPQ-----FQPPAANQLITLSN--LHLS 576  
 QY 531 IVQAAALPGIV-PFVFAKEYGVHPDIL-STAYGPITSHG 567  
 Db 577 LSSLLANNAVNPPTNPQASHPAIGSTTAGSVTISG 615

RESULT 9  
 PNT2\_DROME STANDARD; PRT; 718 AA.  
 ID PNT2\_DROME  
 AC P51023; Q8IG92; Q9VCN2;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE ETS-like protein pointed, isoform P2 (D-ETS-2).  
 GN PNT OR ETS58AB OR ETS2 OR CG17077.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.  
 RX MEDLINE=94038653; PubMed=8223245;  
 RA Klaumbt C.;  
 RT "The Drosophila gene pointed encodes two ETS-like proteins which are involved in the development of the midline glial cells.";  
 RL Development 117:163-176(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Gelniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.O., Scher S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,



RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,  
 RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
 RA Burtis K.C., Buesam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Chetty J.M., Casley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos K., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischer W.,  
 RA Fowler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaeser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D.A., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ileguam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.S., Kodira C.B., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laiko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders A.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Klamis I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RT "The genome sequence of *Drosophila melanogaster*."  
 RL Science 287:2185-2195 (2000).  
 RN [3]  
 RP SEQUENCE OF 1-323 FROM N.A.  
 RC STRAIN=Berkeley; TISSUE=Embryo;  
 EX MEDLINE=22426066; PubMed=12537569;  
 RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,  
 RA George R.A., Guarini H., Krommiller B., Pacleb J.M., Park S., Wan K.H.,  
 RA Rubin G.M., Celiker S.E.;  
 RT "A *Drosophila* full-length cDNA resource."  
 RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).  
 RN [4]  
 RP SEQUENCE OF 551-708 FROM N.A., TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.  
 RC STRAIN=Canton-S; TISSUE=Larva;  
 EX MEDLINE=92249640; PubMed=1571786;  
 RA Chen T., Bunting M., Karim F.D., Thummel C.S.;  
 RT "Isolation and characterization of five *Drosophila* genes that encode an ets-related DNA binding domain."  
 RL Dev. Biol. 151:176-191(1992).  
 CC -!- FUNCTION: Required for glial-neuronal cell interactions at the ventral midline which are necessary for the proper elaboration of commissures in the embryonic CNS.  
 CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing; Named isoforms=2;  
 CC Name=P2; Synonyms=B;  
 CC IsoId=P51023-1; Sequence=Displayed;  
 CC Name=P1; Synonyms=C;  
 CC IsoId=P51022-1; Sequence=External;  
 CC -!- TISSUE SPECIFICITY: Expressed in a complex dynamic pattern in early embryos, including the midline and midline glial cells.  
 CC -!- DEVELOPMENTAL STAGE: Expressed throughout development with lower levels during larval development.  
 CC -!- SIMILARITY: Belongs to the ETS family.  
 CC -!- SIMILARITY: Contains 1 pointed (PNT) domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X69167; CAA48917.1; --  
 DR EMBL; AR003742; AAP56125.1; --  
 DR EMBL; BT001893; AAN71862.1; ALT\_SEQ.  
 DR EMBL; M89472; AAC34200.1; --  
 DR PIR; S33168; S33168.  
 DR HSSP; P14921; 2STT.  
 DR FLYBase; FBgn00031118; pnt.  
 DR GO; GO:0005026; P:Type II transforming growth factor-beta rec. .; NAS.  
 DR GO; GO:0006916; P:anti-apoptosis; IGI.  
 DR GO; GO:0007173; P:EGF receptor signaling pathway; NAS.  
 DR GO; GO:0007162; P:terminal region determination; IGI.  
 DR GO; GO:0007179; P:EGF receptor signaling pathway; NAS.  
 DR GO; GO:0007424; P:corro signaling pathway; IGI.  
 DR GO; GO:0007424; P:corro signaling pathway; IGI.  
 DR InterPro; IPR000418; Ets.  
 DR InterPro; IPR002341; HSF ETS.  
 DR InterPro; IPR003118; SAM\_PNT.  
 DR Pfam; PF00178; Ets; 1.  
 DR Pfam; PF02198; SAM\_PNT; 1.  
 DR PRINTS; PR00454; ETSDOMAIN.  
 DR SMART; SM00413; ETS; 1.  
 DR SMART; SM00251; SAM\_PNT; 1.  
 DR PROSITE; PS00345; ETS DOMAIN 1; 1.  
 DR PROSITE; PS00346; ETS DOMAIN 2; 1.  
 DR PROSITE; PS00061; ETS DOMAIN 3; 1.  
 KW DNA-binding; Nuclear protein; Developmental protein;  
 KW Alternative splicing;  
 FT DOMAIN 166 250 POINTED.  
 FT DOMAIN 341 347 POLY-GLN.  
 FT DOMAIN 359 362 POLY-ASN.  
 FT DOMAIN 389 393 POLY-ASN.  
 FT DOMAIN 405 415 POLY-ASN.  
 FT DOMAIN 418 421 POLY-ALA.  
 FT DOMAIN 481 484 POLY-GLY.  
 FT DNA BIND 610 690 ETS-DOMAIN.  
 FT CONFLICT 133 135 DIS -> YVP (IN REF. 1).  
 SQ SEQUENCE 718 AA; 77683 MW; FDGAFD0F4ECD89C5 CRC64;  
 Query Match 4.1%; Score 119; DB 1; Length 718;  
 Best Local Similarity 20.6%; Pred. No. 0.91;  
 Matches 69; Conservative 50; Mismatches 114; Indels 102; Gaps 13;  
 Qy 197 EAEPDGVAGAVSSRGCGDAGRVTVRKSTSSRSEACSHSOTMOPRVSNLGVVEIY 256  
 Db 344 QSQPTVNGSGGASSNNNS-MLPFAVQSSNNNTSSNNNSNN---NNNSGGNN 399  
 Qy 257 SLOSSRNPTRGSSFNHADFFNIVGAAA-----KGGGGAAGDEKAGCGGGGGHSPQ 310  
 Db 400 SNAGSNN-----NNNNNNINFMAAAATFQHHLKEEPT-----QNGNIGGYGGSSNQ 449  
 Qy 311 QAVAVPAKRKDLHMLVW--SSSASPVSERAAVHVFAGGAD---HADVLAKGAQAYDEVG 365  
 Db 450 DPTDLSYGLPAHLYAAYGGSGSGTGRSS---GGGDSRDYHSTISAQHQSQSSG 505  
 Qy 366 RDDSRTKNGSGGADKGGFTLSKLGNSSTAOLYKPDDEGR----- 407  
 Db 506 -----GNGSGGSGSGTNGSNGYLDSSSEFYSYAGRNRFDGYPPEFTPYDAQSFQ 557  
 Qy 408 -----AAVAMPAPASVMTLIL-----IMVWRKLI 432  
 Db 558 SMCQPQTAMDQGAHAHQHAPARYMSTGLDKLLGGYTTQGGVPCFTGSGFIQLWQLL 617  
 Qy 433 -----RNPNYSSLIQVWVSL-----VSYRWGI 455  
 Db 618 ELLLDKTCQSFISWTGDGWEFKLTPDPEVARRWGI 652  
 RESULT 10  
 MDCI\_RHIME



530 AIQVQALPQGIQVFFFAKEYGV 551  
 :: ||| || ::|||  
 261 AVLLASLPATNPFVIGQQYGV 282

RESULT 11  
 FUS\_MOUSE  
 ID FUS\_MOUSE STANDARD; PRT; 518 AA.  
 AC P56559;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE RNA-binding protein FUS (Pigpen protein).  
 GNGN FUS.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 [1]  
 SEQUENCE FROM N.A.  
 MEDLINE=2829580; PubMed=12950080;  
 RX Alappat S.R.; Zhang M.; Zhao X.; Alliegro M.A.; Alliegro M.C.;  
 RA Burdett C.A.;  
 RT "Mouse pigpen encodes a nuclear protein whose expression is  
 developmentally regulated during craniofacial morphogenesis";  
 Dev. Dyn. 228:59-71 (2003).  
 RL -!- FUNCTION: Binds both single-stranded and double-stranded DNA and  
 promotes ATP-dependent annealing of complementary single-  
 stranded DNAs and D-loop formation in superhelical double-stranded  
 DNA. May play a role in maintenance of genomic integrity (By  
 similarity).  
 CC -!- SUBUNIT: Component of nuclear riboprotein complexes. Interacts  
 with ILF3 and SF1 (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).  
 CC -!- SIMILARITY: Contains 1 RANBP2-type zinc finger.  
 CC -!- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.  
 CC -!- SIMILARITY: BELONGS TO THE TET FAMILY OF RNP PROTEINS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 CC EMBL; AF224264; BAF70602.1; --  
 DR HSSP; P09651; IHA1.  
 DR MGD; MGI:1353633; FUS.  
 DR InterPro; IPR000504; RNA\_rec\_mot.  
 DR InterPro; IPR001876; Znf\_RanGDP.  
 DR Pfam; PF00076; rtm; 1.  
 DR Pfam; PF00641; zf-RanBP; 1.  
 DR SMART; SM00360; RRM; 1.  
 DR SMART; SM00547; Znf\_RBZ; 1.  
 DR PROSITE; PS50102; REM; 1.  
 DR PROSITE; PS00030; RRM\_RNP\_1; FALSE\_NEG.  
 DR PROSITE; PS01358; ZF\_RANBP2\_1; 1.  
 DR PROSITE; PS50199; ZF\_RANBP2\_2; 1.  
 DR RNA-binding; DNA-binding; Nuclear protein; Repeat; Zinc-finger; Zinc;  
 Metal-binding.  
 DR DOMAIN 1 167 GLN/GLY/SER/TYR-RICH.  
 FT DOMAIN 168 265 GLY-RICH.  
 FT DOMAIN 266 364 RNA-BINDING (RRM).  
 FT DOMAIN 364 518 ARG/GLY-RICH.  
 FT ZN FING 415 446 RANBP2-TYPE.  
 SQ SEQUENCE 518 AA; 52673 MW; E06P231BFED78D6 CRC64;

Query Match 4.0%; Score 116; DB 1; Length 518;  
 Best local Similarity 22.5%; Pred. No. 0.99; Indels 94; Gaps 11;  
 Matches 56; Conservative 24; Mismatches 75;

Db 70 QAPQGYSTGYGSS-----QSSQSSYQQSSYGYGQ--QPAFSSYSGSYGG 116  
 QY 257 SIQSRNPTRG-----SSFNHADFF---NIVGAAGKGGGAAG- 292  
 Db 117 SSSSSYGPQSGGYGQSGYGGQSSYQSSYNPPGYGQQNQYNSSSGGGGGGG 176  
 QY 293 ----DEEKACGGGGGSHSPQAVAVPAKRDLEHLVWSSASPVSRAAHVFGAGGA 348  
 Db 177 NYGQDSSMSGGGGGGYGNQDS-----GGGGG 205  
 QY 349 DHA-----DVLAKGAGAYDEYGRDD---YSRTKNGS-----GGADKGPTLSKLGNSST 395  
 Db 206 GYGGGQDGRGRGGGGYNNSSGVEPRGGRGGGGRGGGGSDRGG---FNKFGG--- 260  
 QY 396 AQLYKDDG 404  
 Db 261 ----PRDQG 265

## RESULT 12

KNRL\_DROME  
 ID KNRL\_DROME STANDARD; PRT; 647 AA.  
 AC F13054; Q9VPC8;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Knirps-related protein.  
 GN KNRL OR NR0A2 OR CG4761.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89057149; PubMed=2848202;  
 RA Oro A.E., Ong E.S., Margolis J.S., Posakony J.W., McKeown M.,  
 Evans R.M.;  
 RT "The Drosophila gene knirps-related is a member of the  
 RT steroid-receptor gene superfamily.",  
 RL Nature 336:493-496(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 Ananidis P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.P.,  
 George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,  
 Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 Abril J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,  
 Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 Borkova D., Botchan M.R., Bouck J., Brockstein P., Brottier P.,  
 Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 Foele C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 Glodek A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
 Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 Lasok P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
 Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,  
 Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleeb J.M.,  
 Palazzolo M., Pittman G.S., Fan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 Shue B.C., Siden-Kiamis I., Simpson M., Skupeki M.P., Smith T.,  
 Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,  
 Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RA "The genome sequence of Drosophila melanogaster";  
 RL Science 287:2185-2195(2000).  
 RN [3]  
 RP REVISIONS.  
 RX MEDLINE=22426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,  
 Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
 Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,  
 Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 Lewis S.E.;  
 RA "Annotation of the Drosophila melanogaster euchromatic genome: a  
 RT systematic review";  
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkley; TISSUE=Embryo;  
 RX MEDLINE=22426066; PubMed=12537569;  
 RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,  
 George R.A., Guarin H., Krommiller B., Pacleeb J.M., Park S., Wan K.H.,  
 Rubin G.M., Celniker S.E.;  
 RA "A Drosophila full-length cDNA resource";  
 RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- SIMILARITY: Belongs to the nuclear hormone receptor family. NRO  
 subfamily.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X14153; CAA32365.1; --  
 CC EMBL; AE003591; AAF51627.2; --  
 CC EMBL; AY075383; AAL68221.1; --  
 CC PIR; S06450; S06450.  
 CC HSSP; P20393; 1A6Y.  
 CC TRANSFAC; T02773; --  
 CC FlyBase; FBgn0001323; knr1.  
 CC GO; GO:0004879; F:ligand-dependent nuclear receptor activity; NAS.  
 CC GO; GO:0007088; P:regulation of mitosis; IMP.  
 CC InterPro; IPR001628; Znf\_C4steroid.  
 CC Pfam; PF00105; zf-C4; 1.  
 CC PRINTS; PR00047; STROIDFINGER.  
 CC ProDom; PD000035; Znf\_C4steroid; 1.  
 CC SMART; SM00399; Znf\_C4; 1.  
 CC PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
 CC Receptor; Transcription regulation; DNA-binding; Nuclear protein;  
 CC Zinc-finger; Developmental protein;  
 CC DNA BIND 14 80 NUCLEAR RECEPTOR-TYPE.  
 CC ZN FING 14 34 C4-TYPE.  
 CC ZN FING 51 75 C4-TYPE.  
 CC DOMAIN 121 262 GLY-RICH.  
 CC DOMAIN 577 597 ASN-RICH.  
 CC SEQUENCE 647 AA; 68322 MW; E466FA081DAACDC8 CRC64;

Query Match 4.0%; Score 115; DB 1; Length 647;  
 Best Local Similarity 21.7%; Pred. No. 1.5;  
 Matches 68; Conservative 33; Mismatches 99; Indels 114; Gaps 15;

QY 200 PDGAGAGAYSSRGDAGRVVTVRKSTSRSEAAAC-----SHSHSQTMOPRV 247  
 Db 160 PGYVGLYAVANAG-----SRSKEELWMLGLDGSVEYSHKHPVVASPV 206  
 QY 248 SNL-----SGVIYSLQSRNTPRGSSFNADFFNIVGAAKGGGAGDE--EKGAC 299  
 Db 207 SSPDSHNSDSVSVSV-----RGNPLH-----LGGKNSGGSSGADGSHSGG 252  
 QY 300 GGGGGHSP--QPQAVAVPAKRKDL-----HML 325  
 Db 253 GGGGGVTPGRP-----PQKDKLSPLPLPPGLASMFVMPAPLPPSHLLPGYHPA 307  
 QY 326 VWS-----SASPVSEAAV-----HVFAGAGADHADVLAQAAY-----DEVGRDYS 370  
 Db 308 LYSHQGLLPTPEQQAQAAVAAVQELFNSSGAGQ--FAPGTSPFANHQHKEEDQP 365  
 QY 371 SRKNSGGGADK-----GGPTLSKLSNSTAQILPKDGDGEGRAAAVAMPAPASVMTLILI 425  
 Db 366 APARSPTTHANNHLLTNGGADELTKRFYLDVAVLSQQQSPPPTTKLPPhSKQDYSISA 425  
 QY 426 MVWRKLIRNPTY 439  
 Db 426 LV-----TPNSES 433

RESULT 13  
 ID\_KLP1\_CHLRE STANDARD; PRT; 776 AA.  
 AC P46870;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Kinesin-like protein KLP1.  
 GN KLP1.  
 OS Chlamydomonas reinhardtii.  
 OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
 OC Chlamydomonadaceae; Chlamydomonas.  
 OX NCBI\_TaxID=3055;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=21gr;  
 RX MEDLINE=94266963; PubMed=8207060;  
 RA Bernstein M., Beech P.L., Katz S.G., Rosenbaum J.L.;  
 RT "A new kinesin-like protein (Klp1) localized to a single microtubule  
 of the Chlamydomonas flagellum.";  
 RL J. Cell Biol. 125:1313-1326(1994).  
 CC -!- FUNCTION: May play a role in rotation or twisting of the central  
 pair microtubules of the flagella axoneme.  
 CC -!- SUBCELLULAR LOCATION: BOUND TO THE CENTRAL PAIR MICROTUBULE C2 OR  
 TO PROJECTIONS THAT ARE ATTACHED TO C2.  
 CC -!- SIMILARITY: Belongs to the kinesin-like protein family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC ENBL; X78589; CAA5326.1; -;  
 DR PIR; A53953; A53953.  
 DR HSPSP; P17119; 3KAR.  
 DR InterPro; IPR001752; kinesin\_motor.  
 DR Pfam; PF00225; kinesin; 1.  
 DR PRINTS; PR00380; KINESINHEAVY.  
 DR SMART; SM00129; KISC; 1.  
 DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.  
 DR PROSITE; PS00412; KINESIN MOTOR DOMAIN2; 1.  
 DR PROSITE; PS00067; KINESIN MOTOR DOMAIN; 1.  
 KW Motor protein; Microtubule; ATP-binding; Coiled coil.  
 FT DOMAIN 1 342 KINESIN-MOTOR (BY SIMILARITY).  
 FT DOMAIN 348 371 COILED COIL (POTENTIAL).  
 FT DOMAIN 579 657 COILED COIL (POTENTIAL).  
 FT

FT DOMAIN 658 776 GLOBULAR (POTENTIAL).  
 FT NP\_BIND 91 98 ATP (POTENTIAL).  
 SQ SEQUENCE 776 AA; 33020 MW; 48646509DE8AA71D CRC64;  
 Query'Match 3.9%; Score 114.5; DB 1; Length 776;  
 Best Local Similarity 25.2%; Pred. NO. 2;  
 Matches 65; Conservative 26; Mismatches 94; Indels 73; Gaps 12;  
 QY 195 ELEAE---PDGVAGAGAYS---SRGGDAGRVVTVRKSTSRSEAAACSHSHS-OTMOPRV 247  
 Db 360 ELKAEALMDITLSGKRVSYDDLTDELRELHATCRRLHGEAPEDLPADSMKRVRETF 419  
 QY 248 SNLSGVEI-----YSIQSRNTPRGSSFNADFFNIVGAAKGGGGA-----AGDEK 296  
 Db 420 KALRAVHVAIKADMATOMATLRRATEEGS-----GAAARGGDSAGFSGVGDVL 468  
 QY 297 GACGGGGGHSPPQAVAVPAKRKDLHMLVWSSSASPVSEAAVHVFG-----AGADHAD 352  
 Db 469 RATGGFTVGHAPLD---ARPPVRSSEL-----GSPGAGASGAELGEPSPGGLHAQ 517  
 QY 353 VLAKGAQAYDEYGRDDYSRTKNGSGGADKGGPTLSKLS-----NSTAQILYP 400  
 Db 518 A-----SSHTDAGSNWGD-AGPLSSPGGTRLAGIFGVSGDRNAVFRYK 560  
 QY 401 KDDGERGAAAVAMPASV 418  
 Db 561 VDVGEGRELAASLKAASI 578

RESULT 14  
 ID\_FTSK\_BIFLO STANDARD; PRT; 969 AA.  
 AC Q8G4H3;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE DNA translocase ftsK.  
 GN FTSK OR BL1411.  
 OS Bifidobacterium longum.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;  
 OC Bifidobacteriaceae; Bifidobacterium.  
 OX NCBI\_TaxID=216816;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NCC 2705;  
 RX MEDLINE=22294977; PubMed=12381787;  
 RA Schell M.A., Karmirantzou M., Snel B., Vilanova D., Berger B.,  
 RA Pessi G., Zwahlen M.-C., Desiere F., Bork P., Delley M.,  
 RA Pridmore R.D., Arigoni F.;  
 RT "The genome sequence of Bifidobacterium longum reflects its adaptation  
 to the human gastrointestinal tract.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).  
 CC -!- FUNCTION: DNA motor protein, which is both required to move DNA  
 out of the region of the septum during cell division and for the  
 septum formation. Tracks DNA in an ATP-dependent manner by  
 generating positive supercoils in front of it and negative  
 supercoils behind it (By similarity).  
 CC -!- SUBUNIT: Homohexamer. This suggests the formation of a ring  
 between the two cells at the septum that surrounds DNA (By  
 similarity).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Located at the  
 septum. The large C-terminal part of the protein is cytoplasmic  
 (Potential).  
 CC -!- SIMILARITY: Contains 1 FtsK domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC

```

DR EMBL; AE014771; AAN25210.1; -.
DR HAVAP; MF_01809; -.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR002543; FtsK_SpoIIIE.
DR InterPro; IPR008253; Marvel.
DR Pfam; PF01580; FtsK_SpoIIIE; 1.
DR Pfam; PF01284; MARVEL; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS0901; FTSK; 1.
KW Chromosome partition; Cell division; ATP-binding; DNA-binding;
KW Transmembrane; Complete proteome.
FT TRANSMEM 73 92
FT TRANSMEM 102 124
FT TRANSMEM 136 158
FT TRANSMEM 168 190
FT TRANSMEM 192 214
FT DOMAIN 598 798
FT NP_BIND 615 622
SQ SEQUENCE 969 AA; 102650 MW; 50711A08044CB794 CRC64;
Query Match 3.8%; Score 112; DB 1; Length 969;
Best Local Similarity 20.9%; Pred. No. 4;
Matches 94; Conservative 65; Mismatches 176; Indels 114; Gaps 19;
QY 41 GINFEVALFAVPLSLFHSFISTNPPFAMNLRFLAADTLQKVALALALAS----- 90
DB 72 GLCFLLVLAFLVCASFWFVSGPFGQLLHAIAAGVGLMSVLPVLLAAVAFRLNRSG 131
QY 91 RGLSPRALG-----LDWSI-----TLFSLSTLPNT-----LVNGIPLLRGMYG 129
DB 132 KGSNNPRVVTGWLLWMSICSIIIDVAIAADHTGDTILQSAGLFGFLGSP-----AWG 188
QY 130 ASSAGTLMQVQVLCIWTMLMFLFEY-----RAARALV-----LDQFPDGA 174
DB 189 LSNVFAIIFVW-----GLFSLMTITGTHVTDLPEDARKIAKIQKVPYPMGQETDGS 244
QY 175 ASIVS-FRVEDSDVSLARGVELEAEPDGVAGAVSSRGDA-----GRVVRVTR 224
DB 245 SQPNEVRVGTTLAFADGVPSHDGDD-----GSDNDQAGDARPSLFAFLGRKSKTE 299
QY 225 KSTSRSEA-----ACSHSQTQMPRVNSLGSVEIYLSQSRPTFRGSPFNADFN 278
DB 300 DKTLDKYAADPFDRAASQHGATAETFPVDPMTG-EIIGAR-----TIASSYDGRPHLS 353
QY 279 IVGAAAKGGGGAAG-----DEEKAGCGGGGGHSPQQAQVAVPAKRDOLHMLVWSSAS 332
DB 354 SPAPADADDGDSRTRVITSGQTVAMPGGAVDDP-----WAPSA 395
QY 333 PVSEAAVHFVGCAGDAHADVLAQQAQVDEYGRDDYSRTKNGSGGADKGGPTLSKLS 392
DB 396 QAGTVLAGAGAGAGAAAGAAAGAAATGAYAGAD-ADGSGVGGQGVPTGGQPNATAGN 454
QY 393 NSTAQLYFKDDGEGRAAAVAMPSPASVMT 421
DB 455 DT-----DDANR--PYQLPDLNLLTK 474
RESULT 15
CCG8 HUMAN
ID CCG8 HUMAN STANDARD; PRT; 425 AA.
AC QWXS5; Q9BX70; Q9BY23;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Voltage-dependent calcium channel gamma-8 subunit (Neuronal voltage-
DE gated calcium channel gamma-8 subunit).
GN CACNG8 OR CACNG6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=21601102; PubMed=11738816;
RA Chu P.-J., Robertson H.M., Best P.M.;
RT "Calcium channel gamma subunits provide insights into the evolution of
this gene family.";
RL Gene 280:37-48(2001).
RN [2]
RP SEQUENCE OF 12-426 FROM N.A.
RX MEDLINE=21100909; PubMed=11170751;
RA Burgess D.L., Gefrides L.A., Foreman P.J., Noebels J.L.;
RT "A cluster of three novel Ca(2+) channel gamma subunit genes on
chromosome 19q13.4: evolution and expression profile of the gamma
subunit gene family.";
RL Genomics 71:339-350(2001).
RN [3]
RP SEQUENCE OF 1-204 FROM N.A.
RX TISSUE=cerebellum;
RC Black J.L. III, Kryzer T.J., Lennon V.A.;
RA "Proposed Homo sapiens voltage-gated calcium channel gamma-6
subunit.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Thought to stabilize the calcium channel in an
inactivated (closed) state (By similarity).
CC -1- SUBUNIT: The L-type calcium channel is composed of five subunits:
alpha-1, alpha-2/delta, beta and gamma.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the PMP-22 / EMP / MP20 family. CACNG
subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; AF361354; AAL50049.1; -.
CC EMBL; AF288388; AAK20031.1; -.
CC EMBL; AF234892; AAK15019.1; ALT_INIT.
CC Genew; HGNC:13628; CACNG8.
CC MIM; 606900; -.
DR GO; GO:0005891; C: voltage-gated calcium channel complex; NAS.
DR GO; GO:0005245; F: voltage-gated calcium channel activity; NAS.
DR GO; GO:0006816; P: calcium ion transport; NAS.
DR InterPro; IPR004031; PMP22_Claudin.
DR InterPro; IPR008368; VDCCGAMMA.
DR Pfam; PF00822; PMP22_Claudin; 1.
DR PRINTS; PR01792; VDCCGAMMA.
DR PRINIS; PR01796; VDCCGAMMA8.
KW Ionic channel; Transmembrane; Ion transport; Voltage-gated channel;
KW Calcium channel.
FT TRANSMEM 19 39
FT TRANSMEM 129 149
FT TRANSMEM 158 178
FT TRANSMEM 208 228
FT TRANSMEM 318 338
FT DOMAIN 247 380
FT CONFLICT 1 2
FT CONFLICT 16 16
FT CONFLICT 67 70
FT CONFLICT 362 362
FT CONFLICT 386 386
FT CONFLICT 394 395
FT CONFLICT 409 409
SQ SEQUENCE 425 AA; 43481 MW; E1105BDAF619D3D CRC64;
Query Match 3.8%; Score 110.5; DB 1; Length 425;
Best Local Similarity 23.5%; Pred. No. 1.9;
Matches 57; Conservative 18; Mismatches 71; Indels 97; Gaps 10;
QY 247 VSNLSGVEIYLSQSRNPTPR-----GSSFNEADFFNIGVAAA----- 284

```

Db	171	LSNIIGVIVISANAGEPCPKRDEEKXNHSYGNFYGGLSFILAE---VIGVLAVNIY	227
Qy	285	-----KGGGAAGDEEKAGCGGGGHS-----	307
Db	228	IERSREAHCCSRSDLLKAGGAG-----CGGGGFSAILRLPSYRFRYRRSRSSSR	279
Qy	308	----PQPAVAVPA-----KRDHLMLVWSSSASFVSERAVHVFGAGGADHADVLAKG	357
Db	280	SSEPSRDSAPGGPGPGFASDTISM--YTLRDPKGSVAAGLAGAGGGGGGAVGAFG	337
Qy	358	AOAYDEYGRDDYSSRTKNGSGGA---DKGGPTLSKLGSNSTAQLYPKDDGEGRAAAVAM	413
Db	338	GAAGGAGGG-----GGGGGAGAERDRGGSS-----GFLTLHNAFPKEAGGGVTVTVTR	386
Qy	414	PPA	416
Db	387	PPA	389

Search completed: March 3, 2004, 09:07:26  
Job time : 24 secs



```

Db 118 IPLKGMVG-NSSGLMVQIVVLCIIWYLMFLFYRGARLILGEQFDD-TAGSIIISF 175
QY 181 RVSDVSVL-ARGDVELEAEPDGVAGAGVSRGGDAGRVVTVRKSTSSRSEAAASHSH 239
Db 176 RVSDSIIISLDGRELQTEAEV-----GEDGKLHVTVRKSTSSRSEVFSHMSH 222
QY 240 ---SQTQPRVSNLSGVETIYSLQSSRNPTPRGSSFNHADFFNIVGAAAKGGGGAAG-- 292
Db 223 GLNSGLSITPRPSNLTWAEIYSLQSSRNPTPRASSFNHTDFYSMV-----NKGNASPRHS 277
QY 293 -----DEKGCAGCGGGGHSPOQAVAPD-----KRDHLHMLV 326
Db 278 NFTNLQFDESG--GLGVGNVPRANGSAYPAPNAGIFSPGGKKXANGKDLHMFV 335
QY 327 WSSSASPVSEAAVHVFAGG-----ADHADVLAKGAQAYDEYGRDDYSRTKNGSG 378
Db 336 WSSSASPVSE--GGLHVFKGDDYGNLDGVAHKK-----YDEGRDEFSGNRPGN 386
QY 379 GADKGGTTLKSLGNSSTAQIYKDDGEGRAAAMVPPASVMTLILIMVWKLIRNPNTY 438
Db 387 GVDKDGVLKSLGSSSTAEHLHPSAANGAEAKFTVMPFTSVMTLILIMVWKLIRNPNTY 446
QY 439 SSLIGTVVSVLSVYRWGIEMPAIIARSISILSDAGLGWAMFSLGLFMALOPRIIACGNKLA 498
Db 447 SSLIGTWSLVSFKNMTPALLIANSIILSNAGLGWAMFSLGLFMALOPRIIACGNKLA 506
QY 499 AIAKGVFVAGPAVMAAASAVGLRGVLLHIAIVQAAALPGQIVPFVFAKEYGVHDPDILST 558
Db 507 SFANGVRFLGPAVMAAASAVGLRGDLRIAIQAAALPGQIVPFVFAKEYGVHDPDILST 566
QY 559 A--YG-----PIT 564
Db 567 GVIFGMLIALPIT 579

```

## RESULT 2

```

ID O81215 PRELIMINARY; PRT; 595 AA.
AC O81215;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Auxin transport protein REH1.
GN REH1.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoideae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98344010; PubMed=9679062;
RA Lueschning C., Gaxiola R.A., Grisafi P., Fink G.R.;
RT "ERL1, a root-specific protein involved in auxin transport, is
RT required for gravitropism in arabidopsis thaliana.";
RL Genes Dev. 12:2175-2187(1998).
DR EMBL; AF056027; AAC39514.1; -.
DR PIR; T02876; T02876.
DR Gramene; O81215; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR InterPro; IPR004776; Auxin_eff.
DR Pfam; PF03547; Auxin_eff; 1.
DR TIGRFAMs; TIGR00946; 2a69; 1.
SQ SEQUENCE 595 AA; 64717 MW; 81703D7382CEC2A8 CRC64;

```

```

Query Match 57.2%; Score 1665; DB 10; Length 595;
Best Local Similarity 58.9%; Pred No. 3 2e-106;
Matches 365; Conservative 54; Mismatches 111; Indels 90; Gaps 17;
QY 1 MITADLYHYLTAVVPLVYVAMTLAGSVRWRIPTDQCSGINRFVAFVPLLSFHTIS 60
Db 1 MITAADFYHVTAMVPLVYVAMTLAGSVKWRIFTDQCSGINRFVAFVPLLSFHTIS 60

```

```

QY 61 TNDPFFAMNRLPLAADTLQKVAIVALLA-----LASGLSSPRALGLDWSITLFSSTLPTNT 116
Db 61 TNNPYTMNIRFATAADTLQKLMVLAAMLTAWSHLISRG-----SLEWTTITLFSSTLPTNT 113
QY 117 LVNGIPELLRGMTGASSAGTLMVQVVLQCIITWYLMFLFYRAARALVLDQFPDGAAS 176
Db 114 LVNGIPELLKGMTGERS--GSLMVQIVVLCIIWYLMFLFMFYRGARMLITQFPD--TAAN 171
QY 177 IUSFRVDSVSL--ARGDVELEAEPDGVAGAGVSRGGDAGRVVTVRKSTSSRSEAA 234
Db 172 IASIVVDPDPVSLDGRDAIETETVK-----EDGRHIVTVRRSNASRSDIY 218
QY 235 CSHSHS-QTQWQPRVSNLSGVETIYSLQSSRNPTPRGSSFNHADFFNIVGAAAKGGGGAAGD 293
Db 219 SRRSMGFSTTTPRPSNLTWAEIYSLQSSRNPTPRGSSFNHTDFYSMVGRSSNFGAADA 278
QY 294 BEKGAC-----GGGGGH--SPQQAIVAP-----AKR 319
Db 279 VRTGATPRPSNTEDDASKPKYPLPASNAAPMAGHYPPAPNPVAVSSAPKGAKAATNGQAKG 338
QY 320 KDLHMLVWSSSSSPVSEAAVHVFAGGADHADVLA-----KGAQAYDIF-GRDDYSS 371
Db 339 EDLHMFVWSSSSSPVSD-----VFGGAPDYNDAAAVKSPKMDGAKDREDYVERDDFSF 393
QY 372 RTKNGSGGADKGGPTLSKLGNSSTAQIYKDDGEGRAAAMVPPASVMTLILIMVWKL 431
Db 394 GNR---GVMDRD-----AEAGDEKAAAGADPDKAMAAFTAMPPTSVMTLILIMVWKL 446
QY 432 IRNPNTYSSLIGVWVSVYRWGIEMPAIIARSISILSDAGLGWAMFSLGLFMALOPRII 491
Db 447 IRNPNTYSSLIGLIVSLVCFRWNFEMPAIVLKSISILSDAGLGWAMFSLGLFMALOPHII 506
QY 492 ACNKLAATAMGVRFVAGPAVMAAASAVGLRGVLLHIAIVQAAALPGQIVPFVFAKEYGV 551
Db 507 ACNKRVATYAMVRFVAGPAVMAAASAVGLRGTLHVAIVQAAALPGQIVPFVFAKEYSV 566
QY 552 HPDILSTA--YG-----PIT 564
Db 567 HPSILSTAVIFGMLIALPIT 586

```

## RESULT 3

```

QY 61 TNDPFFAMNRLPLAADTLQKVAIVALLA-----LASGLSSPRALGLDWSITLFSSTLPTNT 116
Db 61 TNNPYTMNIRFATAADTLQKLMVLAAMLTAWSHLISRG-----SLEWTTITLFSSTLPTNT 113
QY 117 LVNGIPELLRGMTGASSAGTLMVQVVLQCIITWYLMFLFYRAARALVLDQFPDGAAS 176
Db 114 LVNGIPELLKGMTGERS--GSLMVQIVVLCIIWYLMFLFMFYRGARMLITQFPD--TAAN 171
QY 177 IUSFRVDSVSL--ARGDVELEAEPDGVAGAGVSRGGDAGRVVTVRKSTSSRSEAA 234
Db 172 IASIVVDPDPVSLDGRDAIETETVK-----EDGRHIVTVRRSNASRSDIY 218
QY 235 CSHSHS-QTQWQPRVSNLSGVETIYSLQSSRNPTPRGSSFNHADFFNIVGAAAKGGGGAAGD 293
Db 219 SRRSMGFSTTTPRPSNLTWAEIYSLQSSRNPTPRGSSFNHTDFYSMVGRSSNFGAADA 278
QY 294 BEKGAC-----GGGGGH--SPQQAIVAP-----AKR 319
Db 279 VRTGATPRPSNTEDDASKPKYPLPASNAAPMAGHYPPAPNPVAVSSAPKGAKAATNGQAKG 338
QY 320 KDLHMLVWSSSSSPVSEAAVHVFAGGADHADVLA-----KGAQAYDIF-GRDDYSS 371
Db 339 EDLHMFVWSSSSSPVSD-----VFGGAPDYNDAAAVKSPKMDGAKDREDYVERDDFSF 393
QY 372 RTKNGSGGADKGGPTLSKLGNSSTAQIYKDDGEGRAAAMVPPASVMTLILIMVWKL 431
Db 394 GNR---GVMDRD-----AEAGDEKAAAGADPDKAMAAFTAMPPTSVMTLILIMVWKL 446
QY 432 IRNPNTYSSLIGVWVSVYRWGIEMPAIIARSISILSDAGLGWAMFSLGLFMALOPRII 491
Db 447 IRNPNTYSSLIGLIVSLVCFRWNFEMPAIVLKSISILSDAGLGWAMFSLGLFMALOPHII 506
QY 492 ACNKLAATAMGVRFVAGPAVMAAASAVGLRGVLLHIAIVQAAALPGQIVPFVFAKEYGV 551
Db 507 ACNKRVATYAMVRFVAGPAVMAAASAVGLRGTLHVAIVQAAALPGQIVPFVFAKEYSV 566
QY 552 HPDILSTA--YG-----PIT 564
Db 567 HPSILSTAVIFGMLIALPIT 586

```



RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,  
 RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,  
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;  
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis  
 RT thaliana";  
 RL Nature 408:816-820(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Shinn P., Chen H., Cheuk R., Kim C.J., Meyers M.C., Banh J.,  
 RA Bower L., Carninci P., Chung M.K., Goldsmith A.D., Hayaishizaki Y.,  
 RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,  
 RA Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M., Nguyen M.,  
 RA Palm C.J., Pham P.K., Quach H.L., Sakano H., Sakurai T., Satou M.,  
 RA Seki M., Southwick A., Toriumi M., Yamada K., Yu G., Shinozaki K.,  
 RA Davis R.W., Theologis A., Ecker J.R.;  
 RT "Arabidopsis cDNA clones";  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Cheuk R., Chen H., Kim C.J., Meyers M.C., Shinn P., Banh J.,  
 RA Bower L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,  
 RA Hayaishizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,  
 RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,  
 RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,  
 RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,  
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,  
 RA Ecker J.R.;  
 RT "Arabidopsis ORF clones";  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC079676; AAG51807.1; -;  
 DR EMBL; AF372950; AAK50090.1; -;  
 DR EMBL; AY093960; AAM16221.1; -;  
 DR EMBL; G96762; G96762;  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR InterPro; IPR004776; Auxin\_eff.  
 DR Pfam; PF03547; Auxin\_eff; 1.  
 DR TIGRFAMs; TIGR00946; 2a69; 1.  
 SQ SEQUENCE 622 AA; 67018 MW; 333E90DB12F62C83 CRC64;

Query Match 56.6%; Score 1647.5; DB 10; Length 622;  
 Best Local Similarity 56.2%; Pred. No. 5.4e-105;  
 Matches 366; Conservative 56; Mismatches 104; Indels 125; Gaps 18;

QY 1 MITADLYHVLTAIVPLVYVAMTAYGSRVWRRIPTDQCNGINRFVAFVPLLSFHFI 60  
 DB 1 MITADLYHVLTAIVPLVYVAMTAYGSRVWRRIPTDQCNGINRFVAFVPLLSFHFI 60

QY 61 TNDPFAWNLRLAADTLQKVAVALALASGLSPRALGLDWSITLPSLTLPNT 120  
 DB 61 ANNPYANNLRLAADTLQKVAVALALASGLSPRALGLDWSITLPSLTLPNT 117

QY 121 IPLLGMYGASSAGTLMVQVVLQCIITWTMLFLFEYRAARALVLDQFPDGAASIVSF 180  
 DB 118 IPLLGMYG-NFSGDLVQIVVLQCIITWTMLFLFEYRAARALVLDQFPDGAASIVSF 175

QY 181 RYDSDVLSI-ARGDVELEAPDGVAGAVSSRGDAGRVTVRKSTSSRSEACSHS 239  
 DB 176 HVDSIDMSLDGQHPLETEAIK-----EDGKLHVTVRNKRSDIYSRSQ 222

QY 240 SQTMQPRVNLSGVEIYSIQSSRNTPRGSFNHADFNIY----- 280  
 DB 223 GLSATPRPNLTAETIYSIQSSRNTPRGSFNHADFNIY----- 282

QY 281 -----GAAAK-GGGGAG-----DEKGACGGGGGSHSPOP----- 310  
 DB 283 KGTPRPSNVEEDGGPAKTAAGTAAGRFRHYQSGGGGGGGGHAHPAPNCPMSPNTGG 342

QY 311 -----QAVAVPAKR-----KDLHMLVWSSASPVSERAAVHVFAGGADHADVLA 357  
 DB 343 GGGTAAGKNAFVVGKRODNGEDLHMFVWSSASPVS-----VFGGGGGNH-----ADY 394

QY 358 AQAYDEYGRD-----DYSSRTKNGSGADKGGPTLSKLG-----SNSTAQLYP 400  
 DB 358 AQAYDEYGRD-----DYSSRTKNGSGADKGGPTLSKLG-----SNSTAQLYP 400

Db 395 STATNDHQDKVSKISVPOGNSNDNQVEREEFSGNKDDDSKVLATDGNINISNKTTO--- 451  
 QY 401 KDDGEGRAAAVAMPASVMTLLIMVWRKLRINPNYSSLLIGVWVSLVSRWGIEMPAI 460  
 Db 452 -----AKVMPPTSVMTLLIMVWRKLRINPNYSSLLIGVWVSLVSRWGIEMPAI 502  
 QY 461 IARSISILSDAGLGMFSLGLFHALQPRRIACGNKLAATAMGVRFVAGFAVMAASIAV 520  
 Db 503 IAKSISILSDAGLGMFSLGLFHALQPRRIACGNKLAATAMGVRFVAGFAVMAASIAV 562  
 QY 521 GLRGVLLHIAIQAALPQGIPIVFFFAKEYGVHPDILSTA--YG-----PIT 564  
 Db 563 GLRGVLLHIAIQAALPQGIPIVFFFAKEYGVHPDILSTA--YG-----PIT 613

RESULT 4  
 Q8H0E0 PRELIMINARY; PRT; 617 AA.  
 ID Q8H0E0  
 AC Q8H0E0;  
 DT 01-MAR-2003 (TREMELrel. 23, Created)  
 DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)  
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)  
 DE PINI-like auxin transport protein.  
 GN CS-PINI.  
 OS Cucumis sativus (Cucumber).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.  
 OX NCBI\_TaxID=3659;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kamada M., Yamasaki S., Fujii N., Higashitani A., Takahashi H.;  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Motohshi K., Seiji Y., Nobuharu F., Atsushi H., Hideyuki H.;  
 RT "Gravity-induced modification of auxin transport and distribution for  
 RT peg formation in cucumber seedlings: Possible roles for CS-AUX1 and  
 RT CS-PINI";  
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB089897; BAC41319.1; -;  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR InterPro; IPR004776; Auxin\_eff.  
 DR Pfam; PF03547; Auxin\_eff; 1.  
 DR TIGRFAMs; TIGR00946; 2a69; 1.  
 SQ SEQUENCE 617 AA; 67344 MW; D61D6E421F231382 CRC64;

Query Match 56.6%; Score 1647; DB 10; Length 617;  
 Best Local Similarity 56.6%; Pred. No. 5.8e-105;  
 Matches 368; Conservative 54; Mismatches 100; Indels 128; Gaps 21;

QY 1 MITADLYHVLTAIVPLVYVAMTAYGSRVWRRIPTDQCNGINRFVAFVPLLSFHFI 60  
 DB 1 MITADLYHVLTAIVPLVYVAMTAYGSRVWRRIPTDQCNGINRFVAFVPLLSFHFI 60

QY 61 TNDPFAWNLRLAADTLQKVAVALALASGLSPRALGLDWSITLPSLTLPNT 116  
 DB 61 TNDPFAWNLRLAADTLQKVAVALALASGLSPRALGLDWSITLPSLTLPNT 113

QY 117 ILMGIPLLRGMYGASSAGTLMVQVVLQCIITWTMLFLFEYRAARALVLDQFPDGAAS 176  
 DB 114 ILMGIPLLRGMYGASSAGTLMVQVVLQCIITWTMLFLFEYRAARALVLDQFPDGAAS 171

QY 177 IVSFVDSVLSI-ARGDVELEAPDGVAGAVSSRGDAGRVTVRKSTSSRSEAC 235  
 DB 172 IVSFVDSVLSI-ARGDVELEAPDGVAGAVSSRGDAGRVTVRKSTSSRSEAC 218

QY 236 SHSHS-QTMQPRVNLSGVEIYSIQSSRNTPRGSFNHADFNIYVGAAGGGGAGD- 293  
 DB 219 RRSVGLSTTPRPNLTAETIYSIQSSRNTPRGSFNHADFNIYVGAAGGGGAGD- 278

QY 294 -----REKGACGGGG-----CHSP----- 308  
 DB 294 -----REKGACGGGG-----CHSP----- 308

```

Db 279 YGLSASRGTPRPSNYEEEG--GGGGKPRPHYNATTCGNANANANANVHYPAEPNGMF 336
Qy 309 -----QPAVAVPAKPK-----DLHMLVWSSASPVSEAAVHVFGAGADHADV- 353
Db 337 STGSKNAQPNNAKPAKGTEDGGDLHMFVWSSASPVSDVFGNHFEGAHN-DQKDV 395
Qy 354 --LAQAQA--YDEYG-RDDYS-----SRTKNGSGGADKGGPTLSKLSGNSTLAQLYPK 401
Db 396 LAVSPGKEGRRENQEBYEREDFSFGNEMMNSNNGGVGVG-----GTEKVGDIKPK 448
Qy 402 DDEGGRAAVAMPASVMTLILIMVWRKLIENPTYSSLIQVWVSLVSYRWGTEMPALI 461
Db 449 -----TMEPTSVMTLILIMVWRKLIENPTYSSLIQVWVSLVSYRWGTEMPALI 498
Qy 462 ARSISILSDAGLGMAFMFSLGLFMALQPRIIACGNKLAALAMGVRFVAGPAAVMAASIAVG 521
Db 499 AKSISILSDAGLGMAFMFSLGLFMALQPRIIACGNKLAALAMGVRFVAGPAAVMAASIAVG 558
Qy 522 LRGVLLHIAIQAALPQGIQVFPVFAKEYGVHDPDILSTA--YG-----PIT 564
Db 559 LRGVLLHIAIQAALPQGIQVFPVFAKEYGVHDPDILSTGVIFQMLVALPIT 608

RESULT 5
Q9ZSY6 PRELIMINARY; PRT; 622 AA.
AC Q9ZSY6;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative auxin efflux carrier protein.
GN PIN1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=99074368; PubMed=9856939;
RA Gailweller L., Guan C., Mueller A., Wisman E., Mendgen K.,
RA Yephremov A., Palme K.;
RT "Regulation of polar auxin transport by AtPIN1 in Arabidopsis vascular
RL tissue."
RL Science 282:2226-2230 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Gailweller L., Changhui G., Mueller A., Wisman E., Palme K.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF089084; AAD04376.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR InterPro; IPR004776; Auxin_eff.
DR Pfam; PF03547; Auxin_eff; I.
DR TIGRFAMs; TIGR00946; 2a69; 1.
SQ SEQUENCE 622 AA; 67031 MW; 8A50F9C2793E8CC5 CRC64;

```

```

Query Match 56.4%; Score 1641.5; DB 10; Length 622;
Best Local Similarity 56.1%; Pred. No. 1.4e-104;
Matches 365; Conservative 56; Mismatches 105; Indels 125; Gaps 18;

Qy 1 MITALDLYHYLTAVVPLVYAMTAYGSRVWRRIPTDQCNGINRFVALFAVPLLSFHFIS 60
Db 1 MITAADFYHYMTAVVPLVYAMTAYGSRVWRRIPTDQCNGINRFVALFAVPLLSFHFIA 60
Qy 61 TNDPFANLRLFLAADTLQKVAVALALASRGLSSPRALGLDWSITLFSLSLTPNLVNG 120
Db 61 ANNPFANLRLFLAADSLQKIVIVLSLLFLWCK-LS--RNGSLDWTITLFSLSLTPNLVNG 117
Qy 121 IPLLGRMYGASSAGTLMVQVVLQCIITWYIMLFLFYRAARALVDQPPDGAASIVSF 180
Db 118 IPLLKMYG-NFSGDLNVQIVLQCIITWYIMLFLFYRAARALVDQPPDGAASIVSI 175

```

```

Qy 181 RVSDSVVSL-ARGDVELEAFPCVAGAGAVSRGGDAGRVVTVRKSTSRSEAAASHSH 239
Db 176 HVSDIMSLDGRQPLETEAIK-----EDGKLHVTVRRSNASRSDIYSRRSQ 222
Qy 240 SQTQPRVSNLSGVLYSLOSSENPTPRGSSFNHADFNIV----- 280
Db 223 GLSATPRPSNLTNWAIYSLOSSENPTPRGSSFNHADFNIV----- 282
Qy 281 -----GAAAK--GGGGAAG--DDEKACCGGGGGHSPQP----- 310
Db 283 KGP:PRPSNYEEDGPAKPTAAGTAAGAGRFHYQSGSGGGGGAHYPAEPNGMFSPNTGG 342
Qy 311 -----QAVAVPAKR-----KOLHMLVWSSASPVSEAAVHVFGAGADHADVLAKG 357
Db 343 GGGTAAGNAAPVVGGRQDGNHDLHMFVWSSASPVSD-----VFGGGGNNH-----ADY 394
Qy 358 AQAYDEYGRD-----DYSRTKNGSGGADKGGPTLSKLSG-----SNSTAQLYP 400
Db 395 STATNDHQKDVKISVPGNSNDNQYVEREEFSGNKDDSKVLATDGNNSNKTTQ--- 451
Qy 401 KDDGGRAAAVAMPASVMTLILIMVWRKLIENPTYSSLIQVWVSLVSYRWGTEMPAI 460
Db 452 -----AKVMPPTSVMTLILIMVWRKLIENPTYSSLIQVWVSLVSYRWGTEMPAI 502
Qy 461 IARSISILSDAGLGMAFMFSLGLFMALQPRIIACGNKLAALAMGVRFVAGPAAVMAASIAV 520
Db 503 IAKSISILSDAGLGMAFMFSLGLFMALQPRIIACGNKLAALAMGVRFVAGPAAVMAASIAV 562
Qy 521 LRGVLLHIAIQAALPQGIQVFPVFAKEYGVHDPDILSTA--YG-----PIT 564
Db 563 LRGVLLHIAIQAALPQGIQVFPVFAKEYGVHDPDILSTGVIFQMLVALPIT 613

RESULT 6
Q9FVF6 PRELIMINARY; PRT; 614 AA.
AC Q9FVF6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE PIN1-like auxin transport protein.
GN PPL1.
OS Populus tremula x Populus tremuloides.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Malpighiales; Salicaceae; Salicaceae; Populus.
OX NCBI_TaxID=47664;
RN [1]
RP SEQUENCE FROM N.A.
RC Schrader J., Bhalerao R.P., Palme K., Sandberg G.;
RA "The PPL family of PIN1-like auxin transporters in Hybrid Aspen.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF190881; AAG17172.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR InterPro; IPR004776; Auxin_eff.
DR Pfam; PF03547; Auxin_eff; I.
DR TIGRFAMs; TIGR00946; 2a69; 1.
SQ SEQUENCE 614 AA; 67254 MW; 7DA7EBD27570ACC9 CRC64;

Query Match 55.5%; Score 1614.5; DB 10; Length 614;
Best Local Similarity 55.6%; Pred. No. 9.9e-103;
Matches 362; Conservative 62; Mismatches 94; Indels 133; Gaps 22;

Qy 1 MITALDLYHYLTAVVPLVYAMTAYGSRVWRRIPTDQCNGINRFVALFAVPLLSFHFIS 60
Db 1 MISUTDFHYMTAVVPLVYAMTAYGSRVWRRIPTDQCNGINRFVALFAVPLLSFHFIS 60
Qy 61 TNDPFANLRLFLAADTLQKVAVALALASRGLSSPRALGLDWSITLFSLSLTPNT 116
Db 61 TNDPFANLRLFLAADSLQKIVIVLSLLFLWCK-LS--RNGSLDWTITLFSLSLTPNT 113
Qy 117 LVNGIPLLRGMYGASSAGTLMVQVVLQCIITWYIMLFLFYRAARALVDQPPDGAAS 176

```

Db 114 LVWGIFLLKMGVGDYS-GSLMVQVVVLOCIITWYLLMFLFEYRAARALVLDQFPDGAASIVSF 171  
QY 177 IYSFRVDSVWSL-ARGDVELAEAPDGVAGAVSSRGEDAGRVRVTVRKSTSSRBAAC 235  
Db 172 IYSIHVDSIMSLDGRPLETEA-----AIK-----EDGKLHVTVRKSNASRSDIFS 218  
QY 236 SHSHS-OTWQPRVSNLSGVEIYSLOSSRNPTPRGSSFNHADFFINVGAAKGGGAG-- 292  
Db 219 RRSQGLSSTTPRPSNLTNDAEIIYSLOSSRNPTPRGSSFNHTDFYSMAAGRNSNFASDVY 278  
QY 293 -----DEKGCAC-----GGGGGHSPOP----- 310  
Db 279 GLSASRGPTPRPSNFFEEHGGGKPRFHHYAPGGAATHYAPAPNPGMFSPTTAASKGVSAN 338  
QY 311 --QAVAVPAK--:::-----KDLHMLVWSSSASPVSERAAVHVFGAGGADHADV-- 353  
Db 339 ANNAARAAKNGQAOQAEDGRDLHMFVWSSSASPVSVDVFGEDYCA--HDLKDVRA 396  
QY 354 LAKG-----AQAYDEYV--RDYS-----STKNGSGGADKGGFTLSKLSNSTAOLYP 400  
Db 397 VSPGKVGORENQEDYNLERDDFSFNGRLDRENSHSG-BKGG----- 439  
QY 401 KDDGEGRAAAVAMPSPASVMTLILIMVWKLIRNNTYSSSLIGVWSSVSVRWGIEMPAI 460  
Db 440 -FDGKPK-----AMPPTSVMTLILIMVWKLIRNNTYSSSLIGLWSLVSPFNWVKNPLI 494  
QY 461 IARSISILSDAGLGWAMFSLGLFMALQPRIIACGNKLAALANGVRFVAGPAAVMAAIAV 520  
Db 495 IAKSISILSDAGLGWAMFSLGLFMALQPRIIACGNSIATFSNAVRFLTGPAINMAAIAV 554  
QY 521 GLRGVLLHAIYVQALPOGIVPFVFAKEVGHVPDILSTA--YG-----PIT 564  
Db 555 GIRGTLHAIYVQALPOGIVPFVFAKEVGHVPDILSTLSTGVFGMLIALPIT 605

## RESULT 7

Q8LKH1 PRELIMINARY; PRT; 640 AA.  
ID Q8LKH1  
AC Q8LKH1;  
DT 01-OCT-2002 (T-EMBLrel. 22, Created)  
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)  
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)  
DE PIN1-like auxin transport protein.  
GN PIN3.  
OS Populus tremula x Populus tremuloides.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids I; Malpighiales; Salicaceae; Salicaceae; Populus.  
OX NCBI\_TaxID=47664;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Baba K., Schrader J., Palme X., Bhalarao R.P., Sandberg G.;  
RT "A family of PIN1 like auxin transporters in hybrid aspen."  
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF515434; AAM54033.1; -  
DR GO; GO:0016021; C-integral to membrane; IEA.  
DR InterPro; IPR004776; Auxin\_eff.  
DR Pfam; PF03547; Auxin\_eff; 1.  
DR TIGRFAMs; TIGR00946; 2a69; 1.  
SQ SEQUENCE 640 AA; 69528 MW; 71C8F4C7C9D40D40 CRC64;

Query Match 55.2%; Score 1605.5; DB 10; Length 640;  
Best Local Similarity 54.7%; Pred. No. 4.4e-102;  
Matches 359; Conservative 71; Mismatches 109; Indels 117; Gaps 19;

QY 1 MITALDLVHLTVAVPLVYAMTILAYGSRVWRIFTPDQCSGINRFVALFAVPLLSHFHIS 60  
Db 1 MTSNDLNVLSAVIPLYVAMTILAYGSRVWRKIFSPDQCSGINRFVAIFAVALLSHFHIS 60  
QY 61 TNDPFAMLRFLAADTLQKVAVALALLASRLGSSPRALGDSITLFSLSLTPNTLVMG 120  
Db 61 TNDPYAMNFRFAADTLQKIIMLIALGIWNTFTKNG---SLEWMTIFSSTLTPNTLVMG 117

QY 121 IPLLGMYGASSAGTLMVQVVVLOCIITWYLLMFLFEYRAARALVLDQFPDGAASIVSF 180  
Db 118 IPLLMTAMTGTYS-GSLMVQIVVLOCIITWYLLMFLFEYRGAKMLIMEQFPE-TAASIVSF 175  
QY 181 RYDSDVWSL-ARGDVELAEAPDGVAGAVSSRGDAGRVRVTVRKSTSSRBAACSHSH 239  
Db 176 KYDSVSVSLDGRDPLETDAEI-----GDGKLVTVRKSNASR--RSLGPGS 220  
QY 240 SOTMOPRVSNLSGVEIYSLOSSRNPTPRGSSFNHADFFINVG----- 282  
Db 221 FSGMTPRPSNTJGAEIYLSLOSSRNPTPRGSFNPNDFYSMMVGQFPGRHNSLGPADLYS 280  
QY 283 --AAKGGGAAGDEBKACGGGGGHSPO-----POAVAVP-----AKR 319  
Db 281 VQSSRGPTPRPSNFEEN-CATATLSSPRGFFYPAQVPTSYAPNPEFASTVTTKTAKN 339  
QY 320 -----KDLHMLVWSSSASPVSERAAVHVFGA---GGADHADVLAKGA----- 358  
Db 340 QQQQNSKANHDAKELHMFVWSSSASPVSSEGGLHVFGGTDFGASBOGSRSDQAKETML 399  
QY 359 -----QAYDEYGRDDYSRTKNGSGAD---KGGPT-LSKLGSNSTAOLYP 400  
Db 400 VADHPQNGETKTIPQDGDPAGEDFSFAGREGDDVDQREKEGPTGLNKLSSSTAELOP 459  
QY 401 K-----DDEGEGRAAAVAMPSPASVMTLILIMVWKLIRNNTYSSSLIGVWSSVSVRWGI 455  
Db 460 KAAEAPDGGGR-----KMPSPASVMTLILIMVWKLIRNNTYSSSLIGLWSLVAFRHHV 515  
QY 456 EMPATIAARSISILSDAGLGWAMFSLGLFMALQPRIIACGNKLAALANGVRFVAGPAAVMAA 515  
Db 516 EMPKIKQISILSDAGLGWAMFSLGLFMALQPKLIACGNSVATFAMAVRFLTGPAAVMAA 575  
QY 516 ASIAGVLLHAIYVQALPOGIVPFVFAKEVGHVPDILSTA--YG-----PIT 564  
Db 576 ASIAGVLLHAIYVQALPOGIVPFVFAKEVGHVPDILSTAVIFGMLIALPIT 631

## RESULT 8

Q7X9P6 PRELIMINARY; PRT; 619 AA.  
ID Q7X9P6  
AC Q7X9P6;  
DT 01-OCT-2003 (T-EMBLrel. 25, Created)  
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)  
DE PIN1-like protein.  
OS Populus tomentosa.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids I; Malpighiales; Salicaceae; Salicaceae; Populus.  
OX NCBI\_TaxID=118781;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Fan J.H.;  
RT "Isolation and expression of Populus tomentosa pin1-like gene."  
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY302060; AAP59843.1; -  
SQ SEQUENCE 619 AA; 67192 MW; 225F129221B3E9D1 CRC64;

Query Match 54.9%; Score 1599; DB 10; Length 619;  
Best Local Similarity 53.6%; Pred. No. 1.2e-101;  
Matches 354; Conservative 64; Mismatches 100; Indels 142; Gaps 16;

QY 1 MITALDLVHLTVAVPLVYAMTILAYGSRVWRIFTPDQCSGINRFVALFAVPLLSHFHIS 60  
Db 1 MICWNDLVNLSAVIPLYVAMTILAYGSRVWRKIFSPDQCSGINRFVAIFAVALLSHFHIS 60  
QY 61 TNDPFAMLRFLAADTLQKVAVALALLASRLGSSPRALGDSITLFSLSLTPNTLVMG 120  
Db 61 TNDPYAMNFRFAADTLQKIIMLIALGIWNTFTKNG---SLEWMTIFSSTLTPNTLVMG 117  
QY 121 IPLLGMYGASSAGTLMVQVVVLOCIITWYLLMFLFEYRAARALVLDQFPDGAASIVSF 180



QY	61	TNDP	FAMNLRFL	ADTLQK	VAVLALLA	SRGLSS	PRALGLD	WSITLFL	SITLPTNL	VWG	120																																				
Db	61	TNNP	YAMNRF	TAADTLQK	IIMLAL	ITWNTF	TANG--	SLEW	MITFSL	SLPTNL	VWG 117																																				
QY	121	IFLL	RGMYG	ASSAGT	LWQVVV	LOCII	IWYTLML	FLFYR	AARALV	LDDQPD	GGAASIVSF 180																																				
Db	118	IFLL	IAMYGE	S-GT	LWQVVV	LOCII	IWYTLML	FLFYR	GAKL	IMEQPE	-TAASIIISF 175																																				
QY	181	RVD	SDVSL	-ARG	DVELE	ABDPD	VAGAG	VS	RGGD	AGRV	TVRKTSSR	SEREA	CSHS	239																																	
Db	176	KYDS	WVSL	DGRD	FLET	DAEV----	GDDG	KLHV	HRK	SNASR	-----	-R	214																																		
QY	240	SOTM	OPRVS	NTSG	VIYS	LQSR	NPTPRG	SSFN	HADFN	IVGAA	KGGG	GAAD	BEEKGAC	299																																	
Db	215	SFMW	TPRS	NLTG	AEIYS	LS	SSRNPT	PRGS	NFNHTD	FYS	MMG	VAPR	HSNFG	ANDVTSVQS 274																																	
QY	300	GGGG	GGHSP	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	308																																	
Db	275	NSRG	PTPR	SNPE	NP	TATSAQ	TVNS	PKFG	FTPAQ	NVPA	APAP	NP	PESS	GLSKSVSK 334																																	
QY	309	-----	-----	-Q	PAV----	AVPAK----	RK	D	LHML	VW	SSAS	PVS	ERA	AHVH 342																																	
Db	335	NSQ	HL	SOQ	QTQ	TQVQ	PQT	VTS	NGS	AAAA	KTSH	DAKEL	HMF	VW	SSAS	PVS	EAS	GLOV 394																													
QY	343	FG	AG	DHAD	VL	KAG	QAQ	AY	DE	VG	RD	YSSR	-----	-----	TKNG	SG	GA	DKG- 384																													
Db	395	FSG	AA	DYG	-----	-ASD	GSR	SEQ	QAK	I	RML	VSD	DH	PP	NG	VNTK	MG	ETL	GGE 446																												
QY	385	-----	-----	-PT	L	SKL	-GS	N	TAQ	L	YP	KD	D	G	G	R	A	A	V	M	P	P	AS	V	M	T	R 421																				
Db	447	ELK	F	V	K	E	B	E	L	V	E	B	E	V	K	E	P	A	A	L	N	K	G	S	S	T	A	E	L	H	P	-----	-T	A	T	V	H	M	P	P	A	S	V	M	T	R 500	
QY	422	L	L	I	L	I	W	W	R	K	L	I	R	N	P	T	Y	S	L	I	G	V	W	S	L	<																					

```

RESULT 11
Q84WX3 PRELIMINARY;      PRt;      599 AA.
ID Q84WX3
AC Q84WX3;
DT 01-JUN-2003 (TtEMBLrel. 24, Created)
DT 01-JUN-2003 (TtEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TtEMBLrel. 25, Last annotation update)
DE Auxin efflux carrier protein PIN1.
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.
OX NCBI_TaxId=3886;
[1]
RN SEQUENCE FROM N.A.
RP Chawla R., DeMason D.A.;
RT "PSPIN1, putative auxin efflux protein.";
RL Submitted (JAN-2003) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AY222857; AAC38045.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR InterPro; IPR004776; Auxin_eff.
DR Pfam; PF03547; Auxin_eff; 1.
DR TIGRfams; TIGR00946; 2a69; 1.
SQ SEQUENCE 599 AA; 64555 MW; A835C775E9FEFA2B CRC64;
Query Match 54.5%; Score 1587; DB 10; Length 599;
Best Local Similarity 56.2%; Pred. No. 7.5e-101;

```

```

RESULT 2
Q8GV75 PRELIMINARY; PRT; 621 AA.
ID Q8GV75
AC Q8GV75;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Amino efflux carrier protein.
DE PIN3.
GN Medicago truncatula (Barrel medic).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosoids I; Fabales; fabaceae; Papilionoideae; Trifolieae; Medicago.
OX NCBI_TaxID=3980;
[1]
RN SEQUENCE FROM N.A.
RP Schnabel E.L.; Frugoli J.A.;
RA "Putative auxin import and export carrier proteins of Medicago
RT truncatula."
RT Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AY115838; AAWS5299.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR InterPro; IPR004776; Auxin_eff.
DR Pfam; PF03547; Auxin_eff; 1.
DR TIGRFAMs; TIGR00946; 2a69; 1.
DR SEQUENCE 621 AA; 68271 MW; 05DB66F8CF842ABA CRC64;
SQ

```

Query Match 54.4%; Score 1584; DB 10; Length 621;  
Best Local Similarity 54.9%; Pred. No. 1.3e-100;  
Matches 355; Conservative 63; Mismatches 111; Indels 118; Gaps 17;

QY 1 MITALDLYHVLTAIVPLVAVMTLAYGSRVWRRIPTDQCGINRFVALFAVPLLSPHFIS 60  
DB 1 MITLKDLYHVLTAIVPLVAVMTLAYGSRVWRRIPTDQCGINRFVALFAVPLLSPHFIS 60  
QY 61 TNDPAMNRLFAADTLQKVAVALALAL-ASGLSSPRALGLDMSITLFSLTLPNT 116  
DB 61 TNNPQMMFRFAADTLQKIMLVALSLSLWTFKNG-----NLEWITLFSLTLPNT 113  
QY 117 LVMGIPLLRGMYGASSAGTLMVQVVLQCIITWYLMFLFYRAARALVLDQPDGAAS 176  
DB 114 LVMGIPLLIAMYGDS-GSLMVQVVLQCIITWYLMFLFYRAARALVLDQPDGAAS 171  
QY 177 IVSFRVDSVSL-ARGDVELEAEEDPDGAGAGVSSRGDAGRVVTRKSTSRSEAA 235  
DB 172 IVSFKVDSVMSLDGRDFTDASV-----GDDGKLHVTVRKSNASRRSFM 218  
QY 236 SHSHSQTPQPRVSNLSGVEIYSLQSSRNPTPRGSGFNEHDFNIVGAAKGGGAAGD-- 293  
DB 219 N-----TPRPSNLTAETIYSLSS-----TPRGSFNEHDFNIVGAAKGGGAAGD 267  
QY 294 -----EKGACGGGGGGH----- 306  
DB 268 SVQSSSRGPTPRPSNFEEGASPRFGYPAAQTVPSTYPVNPFEFSSTTKPVKNQNL 327  
QY 307 ---SPQQAIVAPAKKDLHMLVWSSSPVSEAAHVH-----GAGGADHADVLA-- 355  
DB 328 MPQQFQVQLTKGSDKELHMFVWSSSPVSEAGLVNFVNSQSEBEGAEIRVVVADE 387  
QY 356 -----KGAQAYDEYGRDYS-SRTKNGSGGADKGGPTLSKLSNSTAQLPKDGD- 404  
DB 388 HNQGETNKGQKEIGEDFKFNGVKVGGQEGPNGP--NKLGSNSTPDLHPKATGV 445  
QY 405 EGRAAIVAMPASVMTLILIMVWKLIRNPNTYSSLGVVMSLVSYRWGEMPAITARS 464  
DB 446 ADSGVKLMPPASVMTLILIMVWKLIRNPNTYSSLGLVSLVAFVDFVEMPKIEKS 505  
QY 465 ISILSDAGLGMAFSLGLFMALQPIRIACGNKLAIAIMGVFVAGPAPVAAAIAVGLRG 524  
DB 506 ISILSDAGLGMAFSLGLFMALQPIRIACGNVASFAMAIRFLTGPVAAAIAVGLRG 565  
QY 525 VLLHIAVQALPQGVVPRVFAKE{GVHPDILSTR--YG-----PIT 564  
DB 566 TLLHIAVQALPQGVVPRVFAKEYNVHVPAILLSTAVIFGLIALPIT 612

RESULT 13  
Q9S7Z8 PRELIMINARY; PRT; 640 AA.

AC Q9S7Z8;  
DT 01-MAY-2000 (TEMBLrel. 13, Created)  
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)  
DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)  
DE Auxin transport protein (Putative auxin transport protein RBH1).  
GN F15H11.14 OR PIN3 OR AT1G70940.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
CX NCBI\_TaxID=3702;  
RN [1]  
SEQUENCE FROM N.A.  
RA Federapfel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.P.,  
RA Altari H., Araujo R., Huizar L., Rowley D., Buehler E., Dunn P.,  
RA Gonzalez A., Kremenetskaia I., Kim C., Lenz C., Li J., Liu S.,  
RA Luros S., Schwartz J., Shinn P., Toriumi M., Vysotskaia V.S.,  
RA Walker M., Yu G., Ecker J., Theologis A., Davis R.W.,  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

SEQUENCE FROM N.A.  
RP Priml J., Wisniewska J., Palme K.,  
RA "PIN gene family in Arabidopsis thaliana";  
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.

[3]  
SEQUENCE FROM N.A.  
RP Southwick A., Nguyen M., Tripp M., Palm C.J., Jones T., Wu T.,  
RA Carninci P., Chen H., Cheuk R., Chan M.M., Chang C.H., Dale J.M.,  
RA Deng J.M., Hayashizaki Y., Hsuan V.W., Lee J.M., Ishida J., Kamiya A.,  
RA Kawai J., Kwei C.J., Narusaka M., Quach H.L., Sakurai T., Satou M.,  
RA Seki M., Shinn P., Tang C.C., Toroumi M., Wallender E.K., Wong C.,  
RA Wu H.C., Yamada K., Yu G., Yuan S., Shinozaki K., Ecker J.,  
RA Theologis A., Davis R.W.,  
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.

[4]  
SEQUENCE FROM N.A.  
RP Nguyen M., Karlin-Neumann G., Southwick A., Tripp M., Miranda M.,  
RA Palm C.J., Bower L., Jones T., Banh J., Carninci P., Chen H.,  
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,  
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,  
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,  
RA Ecker J., Theologis A., Davis R.W.,  
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC008148; AAD55507.1; -  
DR EMBL; AF087818; AAD52695.1; -  
DR EMBL; AF136327; AAM96993.1; -  
DR EMBL; BT002085; AAN72096.1; -  
DR PIR; G96733; G96733.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR InterPro; IPR004776; Auxin\_eff.  
DR Pfam; PFO3547; Auxin\_eff; 1.  
DR TIGRFAMs; TIGR00946; 2a69; 1.  
SEQUENCE 640 AA; 69465 MW; 256F82C8E1ADAD80 CRC64;

Query Match 53.6%; Score 1560.5; DB 10; Length 640;  
Best Local Similarity 52.8%; Pred. No. 5.4e-99;  
Matches 353; Conservative 73; Mismatches 102; Indels 139; Gaps 20;

QY 1 MITALDLYHVLTAIVPLVAVMTLAYGSRVWRRIPTDQCGINRFVALFAVPLLSPHFIS 60  
DB 1 MISHDLYHVLTAIVPLVAVMTLAYGSRVWRRIPTDQCGINRFVALFAVPLLSPHFIS 60  
QY 61 TNDPAMNRLFAADTLQKVAVALALALASRLGSLSPRALGLDMSITLFSLTLPNTLVNG 120  
DB 61 TNNPAMNRLFAADTLQKIMLSLVLA--NFTRSGLSLEWITLFSLTLPNTLVNG 117  
QY 121 IPLLGMYGASSAGTLMVQVVLQCIITWYLMFLFYRAARALVLDQPDGAASIVSF 180  
DB 118 IPLLGMYGYS-GSLMVQVVLQCIITWYLMFLFYRAARALVLDQPDGAASIVSF 175  
QY 181 RVSDVVSLARGD-VELEAPDGVAGAVSSRGDAGRVVTRKSTSRSEAAASHH 239  
DB 176 KVESDVVSLDGHDFLEDAEI-----GDDGKLHVTVRKSNASR-RSFC--- 217  
QY 240 SQTMQPRVSNLSGVEIYSLQSSRNPTPRGSGFNEHDFNIVG-----A 282  
DB 218 GPNTFPSPNLTAETIYSLST-----TPRGSFNEHDFNIVGSGFNEHDFNIVG 273  
QY 283 AAKG-----GGGAAGDEKKGCGGGGSHSPQ----- 310  
DB 274 SSRGTPRPSNFEEGASPRFGYPGGAGSYAPNPPEFSSTTSTANKSVNKNPKD 333  
QY 311 ----QAVAVPA-----KKDLHMLVWSSSPVSEAAHVHFGAGGADH-----ADVL 354  
DB 334 VTNQQTTLPTGKSKNSHDAKELHMFVWSSNGSPVSDRAGLVFG-GAPNDQGRSDQG 392  
QY 355 AK-----GAQAYDEYGRDDYSRTKNGSGGADKGGPTL 387  
DB 393 AKETRLMVLDPQSHNGETKVAHHPASGDFGEGEQFSFAGKEEAEERPKDAENG-----L 445  
QY 388 SKLGSNSTAQLPKDGGEGAAVA---MPPASVMTLILIMVWKLIRNPNTYSSLIGV 444  
DB 446 NKLAPNSTAALQSK-TGLGGAEASQRKNMPPASVMTLILIMVWKLIRNPNTYSSLIGL 504









GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2004, 09:05:58 ; Search time 23 Seconds  
(without alignments)  
1286.160 Million cell updates/sec

Title: US-10-030-884-14  
Perfect score: 2910  
Sequence: 1 MITALDLYHVLRAVPLVLA.....DILSTAYGPIITSHGFIPTCHS 573

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/prodata/2/iaa/5A COMB.pap.\*  
2: /cgn2\_6/prodata/2/iaa/5B COMB.pap.\*  
3: /cgn2\_6/prodata/2/iaa/6A COMB.pap.\*  
4: /cgn2\_6/prodata/2/iaa/6B COMB.pap.\*  
5: /cgn2\_6/prodata/2/iaa/PCTUS COMB.pap.\*  
6: /cgn2\_6/prodata/2/iaa/backfiles1.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	144	4.9	320	4	US-09-489-039A-13804
2	138.5	4.8	738	3	US-08-864-038A-3
3	132	4.5	3546	4	US-09-679-279-13
4	126.5	4.3	984	4	US-09-612-402B-43
5	126.5	4.3	1012	4	US-09-612-402B-2
6	122.5	4.2	982	4	US-09-556-877-176
7	122.5	4.2	982	4	US-09-620-412C-176
8	122.5	4.2	982	4	US-09-598-419-176
9	122.5	4.2	1006	4	US-09-556-877-190
10	122.5	4.2	1006	4	US-09-620-412C-190
11	122.5	4.2	1006	4	US-09-598-419-190
12	122.5	4.2	1013	4	US-09-612-402B-16
13	121	4.2	505	4	US-09-612-402B-17
14	121	4.2	1501	4	US-09-252-991A-20266
15	119.5	4.1	542	4	US-09-252-991A-20266
16	118	4.1	503	4	US-09-252-991A-20266
17	116.5	4.0	509	4	US-09-252-991A-23643
18	114.5	3.9	582	4	US-09-450-072-81
19	114.5	3.9	582	4	US-09-351-348-81
20	114.5	3.9	1492	4	US-09-410-551B-21
21	114.5	3.9	1488	4	US-09-410-551B-17
22	113.5	3.9	553	4	US-09-252-991A-17984
23	113.5	3.9	729	4	US-09-252-991A-17984
24	113.5	3.9	1013	4	US-09-252-991A-17984
25	113	3.9	613	4	US-09-489-039A-13221
26	113	3.9	1700	4	US-09-252-991A-13221
27	111.5	3.8	302	4	US-09-252-991A-29977

28 110.5 3.8 383 4 US-09-252-991A-24300 Sequence 24300, A  
29 110.5 3.8 748 2 US-08-997-080-154 Sequence 154, App  
30 110.5 3.8 748 2 US-08-997-362-154 Sequence 154, App  
31 110.5 3.8 748 2 US-09-095-855-154 Sequence 154, App  
32 110.5 3.8 748 4 US-09-324-542-154 Sequence 154, App  
33 110.5 3.8 748 4 US-09-205-426-154 Sequence 154, App  
34 109 3.7 1248 3 US-08-726-214-16 Sequence 16, Appl  
35 108 3.7 551 4 US-09-489-039A-9083 Sequence 9083, Ap  
36 107.5 3.7 421 3 US-09-000-040-1 Sequence 1, Appl  
37 107.5 3.7 742 3 US-08-791-115B-5 Sequence 5, Appl  
38 107 3.7 4928 3 US-09-036-987A-5 Sequence 5, Appl  
39 107 3.7 4928 3 US-09-370-700-5 Sequence 5, Appl  
40 107 3.7 4928 4 US-09-603-207-5 Sequence 5, Appl  
41 106.5 3.7 278 3 US-09-260-527-3 Sequence 3, Appl  
42 106.5 3.7 280 3 US-09-260-527-1 Sequence 1, Appl  
43 106.5 3.7 526 4 US-09-252-991A-28761 Sequence 28761, A  
44 106 3.6 317 4 US-09-489-039A-9943 Sequence 9943, Ap  
45 106 3.6 508 4 US-09-252-991A-17606 Sequence 17606, A

## ALIGNMENTS

RESULT 1  
US-09-489-039A-13804  
; Sequence 13804, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; FILE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; PRIOR FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 13804  
; LENGTH: 320  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-13804

Query Match 4.9%; Score 144; DB 4; Length 320;  
Best Local Similarity 28.3%; Pred No. 2.7e-05;  
Matches 43; Conservative 28; Mismatches 63; Indels 18; Gaps 5;  
Qy 404 GEGRAAAVMPASVMTKLILIMVWRKILRNPTYSLLIGVMSLVYRGMPIAAR 463  
Db 160 GENGSGTILMLP-----VLWRS-VKPIVWGPLLGVLSAI---GKMPDILLA 205  
Qy 464 SISILSDAGLGWMPFSLGLFWALQPIIACGNKLAATAMGVRFVAGPAAVAAATVGLR 523  
Db 206 SKPGLATAAALPLTGVLSARKLQI---NALIATSTIVKLVQPIAAGLVMLGLH 262  
Qy 524 GVLHIAIVQAALPGQIVFFVFAKEYGVH-PD 554  
Db 263 GSIAITAILMALAAGFFGVVFGNRFVGQSPD 294

## RESULT 2

US-08-864-038A-3  
; Sequence 3, Application US/08864038A  
; Patent No. 6001592  
; GENERAL INFORMATION:  
; APPLICANT: Kunio NAKASHIMA et al.  
; TITLE OF INVENTION: NOVEL POLYPEPTIDE GENE CDNA, VECTOR  
; TITLE OF INVENTION: CONTAINING SAID CDNA, HOST CELLS TRANSFORMED WITH SAID  
; TITLE OF INVENTION: VECTOR, POLYPEPTIDE PRODUCED THEREBY, METHOD OF PRODUCING  
; TITLE OF INVENTION: SAID POLYPEPTIDE, DNA ENCODING SAID POLYPEPTIDE AND ANTIBODY  
; TITLE OF INVENTION: TO SAID POLYPEPTIDE  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:

ALPQGI VPFVFAKEYGVHPDILSTAYGPITSHGF 568

Db 2222 VAGGSVWAGPEV 2234

# RESULT 4

US-09-612-402B-43  
 ; Sequence 43, Application US/09612402B  
 ; Patent No. 6642023  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jackson, W. James  
 ; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof  
 ; Patent No. 6642023  
 ; FILE REFERENCE: 7969-086-999  
 ; CURRENT APPLICATION NUMBER: US/09/612,402B  
 ; CURRENT FILING DATE: 2000-07-06  
 ; PRIOR APPLICATION NUMBER: 08/942,596  
 ; PRIOR FILING DATE: 1997-10-02  
 ; NUMBER OF SEQ ID NOS: 43  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 43  
 ; LENGTH: 984  
 ; TYPE: PRT  
 ; ORGANISM: Chlamydia sp.  
 ; US-09-612-402B-43

Query Match 4.3%; Score 126.5; DB 4; Length 984;  
 Best Local Similarity 22.3%; Pred. No. 0.0066;  
 Matches 123; Conservative 70; Mismatches 196; Indels 163; Gaps 29;  
 QY 36 PDQCSG--INRFVALPAVPLLSFPHI--STNDPFAMNLFRLAADTLQKV-----80  
 Db 50 PLSCFGLNLSGFTVLRGHSLTPTENIRTSNGAALSN---SAADGLFTIEGPKELSFNSNC 106  
 QY 81 -AVLALL--ALASRGLSSPRALGLDWSITLFSLSLTPNLTVMG-----IPLRGMYGA 130  
 Db 107 NSLLAVLPAAATTKGSGTPTTTPNGHIYSKTDL---LLANNEKFSFYNSLVSGDGA 163  
 QY 131 SSAGTLMVQVVLQCIWYTLMLFLFEYRAARALVLDQPDGAASIV-SFRVSDVVS 189  
 Db 164 IDAKSLTVQGISKLCV-----FQENTAQA-----DGGACQVVTFSAMANEAPI 207  
 QY 190 ARGDVELEAEPPDGAGAGAVSSRGDAGRVVTRKSTSSRSE-AACSHSHSQTMDPRVS 248  
 Db 208 A-----FVANVAGVRGGIAAVDQGGQ-----VSSSTSTEDPVVFSRNTAVEFDG 257  
 QY 249 NLGSGVEIYSIQ-----SSRNPTPRGSSFNHADPFNIVGAAAKG 286  
 Db 258 RVGG-GIYSYGNVAFNLNGKTLFLNNVSPVYIAAKQPT-SGQASNTSNNGDGGAI 315  
 QY 287 GCGAAGDEKX-----ACGGGGGHSPOQAVAPAKRKDLHMLVWSSS 330  
 Db 316 NGAQAGSNNSGVSFDGEGVVFSSNVAAKGG-----AIYAKKLSV-----AN 359  
 QY 331 ASPV-----SERAHVHFGAG-----GADHADVLAKG-----AQAYDEYGRDDYSR 372  
 Db 360 CGPVQFLRNANDGGAIVLGESGELSADYGDIIFDGNLKRKTAKENAADVNGVTVSSQA 419  
 QY 373 TKNGSGG-----ADKGG-----PTLSKLGNSSTAQ-----LYPKDDGEGRAAAVAMP-PA 416  
 Db 420 ISMGSGGKITTLAKAGHQILFNDFPIEMANGNNQPAOSSKLLKINDGEGYTDIVFANGS 479  
 QY 417 SVMTRELLIMVWRKLIRN-----PNTYSSL-----IGVWVSLVSYRMGIEMPAIARS 464  
 Db 480 STLYQNVTIEQGRIVLRERAKLSVNSLQSGSLYMEAGSTWDFVTPPQQPPA--ANQ 537  
 QY 465 ISILSDAGLQMA 476  
 Db 538 LITLSNLHLSLS 549

# RESULT 5

US-09-612-402B-2  
 ; Sequence 2, Application US/09612402B

Patent No. 6642023  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jackson, W. James  
 ; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof  
 ; Patent No. 6642023  
 ; FILE REFERENCE: 7969-086-999  
 ; CURRENT APPLICATION NUMBER: US/09/612,402B  
 ; CURRENT FILING DATE: 2000-07-06  
 ; PRIOR APPLICATION NUMBER: 08/942,596  
 ; PRIOR FILING DATE: 1997-10-02  
 ; NUMBER OF SEQ ID NOS: 43  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 2  
 ; LENGTH: 1012  
 ; TYPE: PRT  
 ; ORGANISM: Chlamydia sp.  
 ; US-09-612-402B-2

Query Match 4.3%; Score 126.5; DB 4; Length 1012;  
 Best Local Similarity 22.3%; Pred. No. 0.0069;  
 Matches 123; Conservative 70; Mismatches 196; Indels 163; Gaps 29;  
 QY 36 PDQCSG--INRFVALPAVPLLSFPHI--STNDPFAMNLFRLAADTLQKV-----80  
 Db 78 PLSCFGLNLSGFTVLRGHSLTPTENIRTSNGAALSN---SAADGLFTIEGPKELSFNSNC 134  
 QY 81 -AVLALL--ALASRGLSSPRALGLDWSITLFSLSLTPNLTVMG-----IPLRGMYGA 130  
 Db 135 NSLLAVLPAAATTKGSGTPTTTPNGHIYSKTDL---LLANNEKFSFYNSLVSGDGA 191  
 QY 131 SSAGTLMVQVVLQCIWYTLMLFLFEYRAARALVLDQPDGAASIV-SFRVSDVVS 189  
 Db 192 IDAKSLTVQGISKLCV-----FQENTAQA-----DGGACQVVTFSAMANEAPI 235  
 QY 190 ARGDVELEAEPPDGAGAGAVSSRGDAGRVVTRKSTSSRSE-AACSHSHSQTMDPRVS 248  
 Db 236 A-----FVANVAGVRGGIAAVDQGGQ-----VSSSTSTEDPVVFSRNTAVEFDG 285  
 QY 249 NLGSGVEIYSIQ-----SSRNPTPRGSSFNHADPFNIVGAAAKG 286  
 Db 286 RVGG-GIYSYGNVAFNLNGKTLFLNNVSPVYIAAKQPT-SGQASNTSNNGDGGAI 343  
 QY 287 GCGAAGDEKX-----ACGGGGGHSPOQAVAPAKRKDLHMLVWSSS 330  
 Db 344 NGAQAGSNNSGVSFDGEGVVFSSNVAAKGG-----AIYAKKLSV-----AN 387  
 QY 331 ASPV-----SERAHVHFGAG-----GADHADVLAKG-----AQAYDEYGRDDYSR 372  
 Db 388 CGPVQFLRNANDGGAIVLGESGELSADYGDIIFDGNLKRKTAKENAADVNGVTVSSQA 447  
 QY 373 TKNGSGG-----ADKGG-----PTLSKLGNSSTAQ-----LYPKDDGEGRAAAVAMP-PA 416  
 Db 448 ISMGSGGKITTLAKAGHQILFNDFPIEMANGNNQPAOSSKLLKINDGEGYTDIVFANGS 507  
 QY 417 SVMTRELLIMVWRKLIRN-----PNTYSSL-----IGVWVSLVSYRMGIEMPAIARS 464  
 Db 508 STLYQNVTIEQGRIVLRERAKLSVNSLQSGSLYMEAGSTWDFVTPPQQPPA--ANQ 565  
 QY 465 ISILSDAGLQMA 476  
 Db 566 LITLSNLHLSLS 577

# RESULT 6

US-09-556-877-176  
 ; Sequence 176, Application US/09556877  
 ; Patent No. 6432916  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Probst, Peter  
 ; APPLICANT: Bhatia, Ajay  
 ; APPLICANT: Skeiky, Yasir  
 ; APPLICANT: Fling, Steve

; APPLICANT: Maisonneuve, Jeff  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
 ; FILE REFERENCE: 210121.469C5  
 ; CURRENT APPLICATION NUMBER: US/09/556.877  
 ; CURRENT FILING DATE: 2000-04-19  
 ; NUMBER OF SEQ ID NOS: 305  
 ; SOFTWARE: FastSeq for Windows Version 3.0/4.0  
 ; SEQ ID NO 176  
 ; LENGTH: 982  
 ; TYPE: PRT  
 ; ORGANISM: Chlamydia  
 ; NAME/KEY: VARIANT  
 ; LOCATION: (1)---(982)  
 ; OTHER INFORMATION: Xaa = Any Amino Acid  
 ; US-09-556-877-176

Query Match 4.2%; Score 122.5; DB 4; Length 982;  
 Best Local Similarity 21.4%; Pred. No. 0.016;  
 Matches 137; Conservative 71; Mismatches 222; Indels 209; Gaps 32;

```

QY 36 PDQCSG--INRFVALFAVPLLSFHFI--STNDPPAMNLRFLAADTLQKV----- 80
Db 48 PLSCFNLGSGFTVLGRGHSILTFENIRTSNGAALSN---SAADGLFTIEGFKELSFNC 104
QY 81 -AVLALL--ALASRGLSPRALGLDWSITLFSLSLTPNLTVMG-----IPLRGMYGA 130
Db 105 NSLLAVLPAATTKGSGQTPTTSTPSNGTYSKTDL---LLNNEKFSFVSNLVSDDGA 161
QY 131 SSAGTLMVQVVLQCLIIWYLMFLFEYRAARALVLDQPDGAAASIV--SRVDSDDVSL 189
Db 162 IDAKSLTVQGISKLCV-----FOENTAQA-----DGGACQVTSFAMANEAPI 205
QY 190 ARGDVELEAEPDGVAGAGVSSRGDAGRVRVTVRKSTSSRSB--AACSHSHSQTWPQVRS 248
Db 206 A-----FVANVAGVRGGIAAVDQGGQ-----VSSSTSTEDPVVSFRNTAVDFDGNVA 255
QY 249 NLSGVEIYSLQ-----SSRNPTRGSSFNHADFFNIVGAAAKG 286
Db 256 RVGG-GIYSGNVAFNLNGKTLFLNNVASPVYIAAKQPT-SGQASNTSNYDGGGAI FCK 313
QY 287 GGAAGDEBK-----ACGGGGGHSPOQAVAVPAKRDHLMLVWSSS 330
Db 314 NGAQAGSNNSGVSFDGEGVVFSSNVAAGKGG-----AIYAKKLSV-----AN 357
QY 331 ASPV-----SRAAVHVFAG-----GADHADVLAKG-----AQAYDEXGRDDYSR 372
Db 358 CGPQVFLNIANDGGAIIYLGESGELSADYGLIIFDGNLKRKTAKENAADVNGVTSSQA 417
QY 373 TNGSGG-----ADKGG-----PTLSKLSNSTAQ-----LYPKDDGEGRAAAVAMPAS 417
Db 418 ISMSGGKITTIRAKAGHQILFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDI VFANGS 477
QY 418 VMTRLILIMWRKLRNPNTYSSLGIVVMSLVSYRWGIEMPAIAR-----SISILSDAG 472
Db 478 -----STLYQNTIEQGRIVLRKAKLSVNSLSQTG 508
QY 473 --LGMAMFSLGLFVALQPRIIACGNKLAALAMGVRFVAGPAMVAASIAVGLRGVLLHA 530
Db 509 GSLYMEAGSTLDFVTQP-----PQPPAANQLITLSN--LHLS 545
QY 531 IVQAAALPQGI V-PFVFAKEYGVHPDIL-STAYGPTITSHG 567
Db 546 LSSLLANNVNTNPTNPAPQDSHPAVIGSTTAGSVTISG 584
  
```

## RESULT 7

US-09-620-412C-176  
 ; Sequence 176, Application US/09620412C  
 ; Patent No. 6448234  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Steven P. Fling

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
 ; FILE REFERENCE: 210121.469C7  
 ; CURRENT APPLICATION NUMBER: US/09/620.412C  
 ; CURRENT FILING DATE: 2000-07-20  
 ; NUMBER OF SEQ ID NOS: 363  
 ; SOFTWARE: FastSeq for Windows Version 3.0/4.0  
 ; SEQ ID NO 176  
 ; LENGTH: 982  
 ; TYPE: PRT  
 ; ORGANISM: Chlamydia  
 ; NAME/KEY: VARIANT  
 ; LOCATION: (1)---(982)  
 ; OTHER INFORMATION: Xaa = Any Amino Acid  
 ; US-09-620-412C-176

Query Match 4.2%; Score 122.5; DB 4; Length 982;  
 Best Local Similarity 21.4%; Pred. No. 0.016;  
 Matches 137; Conservative 71; Mismatches 222; Indels 209; Gaps 32;

```

QY 36 PDQCSG--INRFVALFAVPLLSFHFI--STNDPPAMNLRFLAADTLQKV----- 80
Db 48 PLSCFNLGSGFTVLGRGHSILTFENIRTSNGAALSN---SAADGLFTIEGFKELSFNC 104
QY 81 -AVLALL--ALASRGLSPRALGLDWSITLFSLSLTPNLTVMG-----IPLRGMYGA 130
Db 105 NSLLAVLPAATTKGSGQTPTTSTPSNGTYSKTDL---LLNNEKFSFVSNLVSDDGA 161
QY 131 SSAGTLMVQVVLQCLIIWYLMFLFEYRAARALVLDQPDGAAASIV--SRVDSDDVSL 189
Db 162 IDAKSLTVQGISKLCV-----FOENTAQA-----DGGACQVTSFAMANEAPI 205
QY 190 ARGDVELEAEPDGVAGAGVSSRGDAGRVRVTVRKSTSSRSB--AACSHSHSQTWPQVRS 248
Db 206 A-----FVANVAGVRGGIAAVDQGGQ-----VSSSTSTEDPVVSFRNTAVDFDGNVA 255
QY 249 NLSGVEIYSLQ-----SSRNPTRGSSFNHADFFNIVGAAAKG 286
Db 256 RVGG-GIYSGNVAFNLNGKTLFLNNVASPVYIAAKQPT-SGQASNTSNYDGGGAI FCK 313
QY 287 GGAAGDEBK-----ACGGGGGHSPOQAVAVPAKRDHLMLVWSSS 330
Db 314 NGAQAGSNNSGVSFDGEGVVFSSNVAAGKGG-----AIYAKKLSV-----AN 357
QY 331 ASPV-----SRAAVHVFAG-----GADHADVLAKG-----AQAYDEXGRDDYSR 372
Db 358 CGPQVFLNIANDGGAIIYLGESGELSADYGLIIFDGNLKRKTAKENAADVNGVTSSQA 417
QY 373 TNGSGG-----ADKGG-----PTLSKLSNSTAQ-----LYPKDDGEGRAAAVAMPAS 417
Db 418 ISMSGGKITTIRAKAGHQILFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDI VFANGS 477
QY 418 VMTRLILIMWRKLRNPNTYSSLGIVVMSLVSYRWGIEMPAIAR-----SISILSDAG 472
Db 478 -----STLYQNTIEQGRIVLRKAKLSVNSLSQTG 508
QY 473 --LGMAMFSLGLFVALQPRIIACGNKLAALAMGVRFVAGPAMVAASIAVGLRGVLLHA 530
Db 509 GSLYMEAGSTLDFVTQP-----PQPPAANQLITLSN--LHLS 545
QY 531 IVQAAALPQGI V-PFVFAKEYGVHPDIL-STAYGPTITSHG 567
Db 546 LSSLLANNVNTNPTNPAPQDSHPAVIGSTTAGSVTISG 584
  
```

## RESULT 8

US-09-598-419-176  
 ; Sequence 176, Application US/09598419  
 ; Patent No. 6565856  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Skeiky, Yasir A.W.  
 ; APPLICANT: Scholler, John

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND  
 ; FILE REFERENCE: 210121.469C6  
 ; CURRENT APPLICATION NUMBER: US/09/598.419  
 ; CURRENT FILING DATE: 2000-06-20  
 ; NUMBER OF SEQ ID NOS: 357  
 ; SOFTWARE: FastSeq for Windows Version 3.0/4.0  
 ; SEQ ID NO 176  
 ; LENGTH: 982  
 ; TYPE: PRT  
 ; ORGANISM: Chlamydia  
 ; FEATURE:  
 ; NAME/KEY: VARIANT  
 ; LOCATION: (1)...(982)  
 ; OTHER INFORMATION: Xaa = Any Amino Acid  
 US-09-598-419-176

Query Match 4.2%; Score 122.5; DB 4; Length 982;  
 Best Local Similarity 21.4%; Pred. No. 0.016;  
 Matches 137; Conservative 71; Mismatches 222; Indels 209; Gaps 32;  
 Qy 36 PDQCSG--INRFVALFAVPLLSFHFPI--STNDPFAANLRLFLAADTLQKV----- 80  
 Db 48 PLSCFNLGSGFTVLGRGHSILTFENIRTSNGAALSN---SAADGLFTIEGFKLSFNSC 104  
 Qy 81 -AVLALL--ALASRGLSSPRALGLDWSITLPSLSTPLNTLVMG-----IPLLRGMVGA 130  
 Db 105 NSLLAVLPAAATTKNSQPTTTTSPNGTIYKTDL---LLNNEKFSFYSNLVSGDGA 161  
 Qy 131 SSAGTLMVQVVLQCIWIYTLMLFLFEYRAARALVLDQFPDGAASIV--SPRVSDVVS 189  
 Db 162 IDAKSLTVQGISKLCV-----FOENTAQA-----DGGACQVVTFSAMANEAPI 205  
 Qy 190 ARGDVELEAPDGVAGAGAVSSRGDAGRVRVTRKSTSSRSE--AACSHSHSQTWQPRVS 248  
 Db 206 A-----FVANVAGVRGGGIAAVDQCGQ-----VSSSTSTEDPVVFSRNTAVEFDG 255  
 Qy 249 NLGVEIYSLQ-----SSRNPTRGSSFNHADFFNIVGAAAKG 286  
 Db 256 RVGG--GIYSYGNVAFNLNGKTLFLNNVASPVYIAAKQPT--SQASNTSNNGYDGGAI 313  
 Qy 287 GGGAGDDEKG-----ACGGGGGHSPOQAVAVPAKEDLHMLVWSSS 330  
 Db 314 NGAQAGSNNSGVSFDGEGVFFSSNVAAKGG-----AIYAKKLSV-----AN 357  
 Qy 331 ASPV-----SERAHVHVFAG---GADHADVLAKG-----AQAYDEYGRDDYSR 372  
 Db 358 CGPVQFLRNTANDGGAIYLGESGELSADYGDIIFDGNLKRKTAKENAAADVNGVTSSQA 417  
 Qy 373 TKNGSGG-----ADKGG-----PTLSKLGSNSTAQ---LYPKDDGEGRAAAVAMP 417  
 Db 418 ISMGSGKITTIRAKAGHQILFNDPIEMANGNNQPAQSSKLLKINDGEGYTDIVFANGS 477  
 Qy 418 VMTRLLIMVWRKLIIRNPNTYSSLLIGVWVSLVSYRWGIEMPAIAR-----SISILSDAG 472  
 Db 478 -----STLYQNVITIEQGRIVLREKAKLSVNSLQ 508  
 Qy 473 --LGMAMFSLGLFMALOPRIIACGNKLAIAIANGVRVFPAGPAAVMAASIAVGLRGVLLHIA 530  
 Db 509 GSLYWEAGSLDFVTPQP-----PQPPANQLITLSN---LHLS 545  
 Qy 531 IVQAALPQGIIV--PFVFAKEYGVHPDIL--STAYGPITSHG 567  
 Db 546 LSSLLANNAVTPPTNPAPQDHPAVIGSTTAGSVTISG 584

## RESULT 9

US-09-598-877-190  
 ; Sequence 190, Application US/09556877  
 ; Patent No. 6432916  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Probst, Peter  
 ; APPLICANT: Bhatia, Ajay

## RESULT 10

US-09-620-412C-190  
 ; Sequence 190, Application US/09620412C  
 ; Patent No. 6448234  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Steven P. Fling  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
 ; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION

; APPLICANT: Skeiky, Yasir  
 ; APPLICANT: Fling, Steve  
 ; APPLICANT: Maisonneuve, Jeff  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
 ; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION  
 ; FILE REFERENCE: 210121.469C5  
 ; CURRENT APPLICATION NUMBER: US/09/556,877  
 ; CURRENT FILING DATE: 2000-04-19  
 ; NUMBER OF SEQ ID NOS: 305  
 ; SOFTWARE: FastSeq for Windows Version 3.0/4.0  
 ; SEQ ID NO 190  
 ; LENGTH: 1006  
 ; TYPE: PRT  
 ; ORGANISM: Chlamydia  
 US-09-556-877-190

Query Match 4.2%; Score 122.5; DB 4; Length 1006;  
 Best Local Similarity 21.4%; Pred. No. 0.016;  
 Matches 137; Conservative 71; Mismatches 222; Indels 209; Gaps 32;  
 Qy 36 PDQCSG--INRFVALFAVPLLSFHFPI--STNDPFAANLRLFLAADTLQKV----- 80  
 Db 72 PLSCFNLGSGFTVLGRGHSILTFENIRTSNGAALSN---SAADGLFTIEGFKLSFNSC 128  
 Qy 81 -AVLALL--ALASRGLSSPRALGLDWSITLPSLSTPLNTLVMG-----IPLLRGMVGA 130  
 Db 129 NSLLAVLPAAATTKNSQPTTTTSPNGTIYKTDL---LLNNEKFSFYSNLVSGDGA 185  
 Qy 131 SSAGTLMVQVVLQCIWIYTLMLFLFEYRAARALVLDQFPDGAASIV--SPRVSDVVS 189  
 Db 186 IDAKSLTVQGISKLCV-----FOENTAQA-----DGGACQVVTFSAMANEAPI 229  
 Qy 190 ARGDVELEAPDGVAGAGAVSSRGDAGRVRVTRKSTSSRSE--AACSHSHSQTWQPRVS 248  
 Db 230 A-----FVANVAGVRGGGIAAVDQCGQ-----VSSSTSTEDPVVFSRNTAVEFDG 279  
 Qy 249 NLGVEIYSLQ-----SSRNPTRGSSFNHADFFNIVGAAAKG 286  
 Db 280 RVGG--GIYSYGNVAFNLNGKTLFLNNVASPVYIAAKQPT--SQASNTSNNGYDGGAI 337  
 Qy 287 GGGAGDDEKG-----ACGGGGGHSPOQAVAVPAKEDLHMLVWSSS 330  
 Db 338 NGAQAGSNNSGVSFDGEGVFFSSNVAAKGG-----AIYAKKLSV-----AN 381  
 Qy 331 ASPV-----SERAHVHVFAG---GADHADVLAKG-----AQAYDEYGRDDYSR 372  
 Db 382 CGPVQFLRNTANDGGAIYLGESGELSADYGDIIFDGNLKRKTAKENAAADVNGVTSSQA 441  
 Qy 373 TKNGSGG-----ADKGG-----PTLSKLGSNSTAQ---LYPKDDGEGRAAAVAMP 417  
 Db 442 ISMGSGKITTIRAKAGHQILFNDPIEMANGNNQPAQSSKLLKINDGEGYTDIVFANGS 501  
 Qy 418 VMTRLLIMVWRKLIIRNPNTYSSLLIGVWVSLVSYRWGIEMPAIAR-----SISILSDAG 472  
 Db 502 -----STLYQNVITIEQGRIVLREKAKLSVNSLQ 532  
 Qy 473 --LGMAMFSLGLFMALOPRIIACGNKLAIAIANGVRVFPAGPAAVMAASIAVGLRGVLLHIA 530  
 Db 533 GSLYWEAGSLDFVTPQP-----PQPPANQLITLSN---LHLS 569  
 Qy 531 IVQAALPQGIIV--PFVFAKEYGVHPDIL--STAYGPITSHG 567  
 Db 570 LSSLLANNAVTPPTNPAPQDHPAVIGSTTAGSVTISG 608

```

; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620.412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 190
; LENGTH: 1006
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-620-412C-190

```

```

Query Match      4.2%; Score 122.5; DB 4; Length 1006;
Best Local Similarity 21.4%; Pred. No. 0.016;
Matches 137; Conservative 71; Mismatches 222; Indels 209; Gaps 32;

QY 36 PDQCSG--INRFVALFAVPLLSFHEI--STNDPPAMNLRFLAADTLQKV----- 80
DB 72 PLSCFENLGSFTVLGRGHSITFENIRTSNGAALSN---SAADGLFTIEGFKELSFNC 128
QY 81 -AVLALL--ALASRLSPRALGLDWSITLFSLTPLNTLVMG-----IPLRGMVGA 130
DB 129 NSLLAVLPAATNKGSGQPTTTSTPSNGTIYSKTDL---LLLNNEKFSFVSNLVSGDGA 185
QY 131 SSAGTLMVQVVVLOCIWYTLMLFLFEYRAARALVLDQPDGAAASIV-SFRVDSVVS 189
DB 186 IDAKSLTVQGISKLCV-----FOENTAQA-----DGGACQVVTFSAMANEAPI 229
QY 190 ARGDVLEAEPPGVAGAVSSRGDAGRVRVTRKSTSSRSE--AACSHSHSQTWPQPRVS 248
DB 230 A-----FVANVAGVRGGIAAVDQCGQG-----VSSSTSTEDPVVSFRNTAVEFDG 279
QY 249 NLSGVEIYSLO-----SSRNPTPRGSSFNHADPFNIVGAAAKG 286
DB 280 RVGG--GIYSYGNVAFNLNGKTLFLNNVASPVYIAAKQPT--SGQASNTSNNGYDGGAI 337
QY 287 GCGAAGDEBK-----ACGGGGGGHSPQQA VAVPAKRDHLMLVWSSS 330
DB 338 NGAQAGSNNGSVSFDGEGVVFSSNVAAGKGG-----AIYAKLSV-----AN 381
QY 331 ASPV-----SERAHVHFCAG---GADHADVLAKG-----AAQAYDEYGRDDYSR 372
DB 382 CGVQFLRNANDGAIYLGESGELSADYGIIDFNGLKRTAKENAADVNGVTVSSQA 441
QY 373 TKNSSGG-----ADKGG-----PTLSKLGSNSTAQ---LYPKDDGEGRAAAVAMP 417
DB 442 ISMGSGGKITTRAKAGHQILFNDPIEMANGNNQPAQSSKLLKINDGEGYTDIVFANGS 501
QY 418 VMTRLILIMWRKLIRNPNTYSSLGIVVWSLVSRWGIEMPAIAR-----SISILSDAG 472
DB 502 -----STLYQNVTIEQGRIVLREKAKLSVNSLSQTG 532
QY 473 --LGMAMFSLGLFMALQPRIIACGNKLAIAIANGVRFVAGPAVMAAASIAVGLRGVLLH 530
DB 533 GSLYMEAGSTLDFVTPQP-----PQPPAANQLITLSN--LHLS 569
QY 531 IVQAALPQGIIV-PFVFAKEYGVHPDIL-STAYGPITSHG 567
DB 570 LSSLLANNVTPPTNPPAQDSHPAVIGSTTAGSVTISG 608

```

## RESULT 11

```

US-09-598-419-190
; Sequence 190, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Scheller, John
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C6
; CURRENT APPLICATION NUMBER: US/09/598.419
; CURRENT FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 357

```

```

; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 190
; LENGTH: 1006
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-598-419-190

```

```

Query Match      4.2%; Score 122.5; DB 4; Length 1006;
Best Local Similarity 21.4%; Pred. No. 0.016;
Matches 137; Conservative 71; Mismatches 222; Indels 209; Gaps 32;

QY 36 PDQCSG--INRFVALFAVPLLSFHEI--STNDPPAMNLRFLAADTLQKV----- 80
DB 72 PLSCFENLGSFTVLGRGHSITFENIRTSNGAALSN---SAADGLFTIEGFKELSFNC 128
QY 81 -AVLALL--ALASRLSPRALGLDWSITLFSLTPLNTLVMG-----IPLRGMVGA 130
DB 129 NSLLAVLPAATNKGSGQPTTTSTPSNGTIYSKTDL---LLLNNEKFSFVSNLVSGDGA 185
QY 131 SSAGTLMVQVVVLOCIWYTLMLFLFEYRAARALVLDQPDGAAASIV-SFRVDSVVS 189
DB 186 IDAKSLTVQGISKLCV-----FOENTAQA-----DGGACQVVTFSAMANEAPI 229
QY 190 ARGDVLEAEPPGVAGAVSSRGDAGRVRVTRKSTSSRSE--AACSHSHSQTWPQPRVS 248
DB 230 A-----FVANVAGVRGGIAAVDQCGQG-----VSSSTSTEDPVVSFRNTAVEFDG 279
QY 249 NLSGVEIYSLO-----SSRNPTPRGSSFNHADPFNIVGAAAKG 286
DB 280 RVGG--GIYSYGNVAFNLNGKTLFLNNVASPVYIAAKQPT--SGQASNTSNNGYDGGAI 337
QY 287 GCGAAGDEBK-----ACGGGGGGHSPQQA VAVPAKRDHLMLVWSSS 330
DB 338 NGAQAGSNNGSVSFDGEGVVFSSNVAAGKGG-----AIYAKLSV-----AN 381
QY 331 ASPV-----SERAHVHFCAG---GADHADVLAKG-----AAQAYDEYGRDDYSR 372
DB 382 CGVQFLRNANDGAIYLGESGELSADYGIIDFNGLKRTAKENAADVNGVTVSSQA 441
QY 373 TKNSSGG-----ADKGG-----PTLSKLGSNSTAQ---LYPKDDGEGRAAAVAMP 417
DB 442 ISMGSGGKITTRAKAGHQILFNDPIEMANGNNQPAQSSKLLKINDGEGYTDIVFANGS 501
QY 418 VMTRLILIMWRKLIRNPNTYSSLGIVVWSLVSRWGIEMPAIAR-----SISILSDAG 472
DB 502 -----STLYQNVTIEQGRIVLREKAKLSVNSLSQTG 532
QY 473 --LGMAMFSLGLFMALQPRIIACGNKLAIAIANGVRFVAGPAVMAAASIAVGLRGVLLH 530
DB 533 GSLYMEAGSTLDFVTPQP-----PQPPAANQLITLSN--LHLS 569
QY 531 IVQAALPQGIIV-PFVFAKEYGVHPDIL-STAYGPITSHG 567
DB 570 LSSLLANNVTPPTNPPAQDSHPAVIGSTTAGSVTISG 608

```

## RESULT 12

```

US-09-612-402B-16
; Sequence 16, Application US/09612402B
; Patent No. 6642023
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; PATENT NO. 6642023
; FILE REFERENCE: 7969-086-999
; CURRENT APPLICATION NUMBER: US/09/612.402B
; CURRENT FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16

```



```

; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-09-612-402B-16

Query Match
Best Local Similarity 4.2%; Score 122.5; DB 4; Length 1013;
Matches 123; Conservative 66; Mismatches 199; Indels 191; Gaps 30;

QY 81 AVALLL--ALASGLSPRALGLDMSITLFSITLPTLVMG-----IPLLRGMYGAS 131
Db 136 SLIAVLPAATNNGSOTPTTSPNGTIYSKTDL---LLANNEKFSYSLNLSVSGGGTI 192
QY 132 SAGTLMVQVVVLOCIITMYTLMFLFEYRAARALVLDQFPDGAASIV-SFRVDSVWSLA 190
Db 193 DAKSLTVQGISKLCV-----FQENTAQA-----DGGACQVTSFSAMANEAPIA 236
QY 191 RGDVELEAEPDGVAGAGAVSSRGDAGRVRVTRKSTSSRSE-AACSHSHSQTMPRVSN 249
Db 237 -----FTANVAGVRGGIAAVQDQGG-----VSSSTSTEDPVVFSRNTAVEFDGNVA 286
QY 250 LSGVEIYSLQ-----SSRNPTRGSSFNHADFFNIVGAA-AKG 286
Db 287 VCG-GIYSYGNVAFNLNGKTLFLNNVASPVYIAAEQPT-NGQASNTSDNYGGAIFCKN 344
QY 287 GCGAAGDEEKG-----ACGGGGGCHSPQOAVAVPAKDKLHMLVWSSS 330
Db 345 GAGAAGSNNSGSVFQGGGVVFFSNVAAKGKG-----AIYAKKLSV-----AN 388
QY 331 ASPV-----SERAHVHVFAG-----GADHADVLAKG-----AQAYDEYGRDDYSSR 372
Db 389 CGPVQFLGNANDGGAIVLGESGELSADYDGIIFDGNLKRKTAKENAADVNGVTVSQA 448
QY 373 TKNGSG-----ADKGG-----PTLSKLSNSTAQ-----LYPKDDGEG 406
Db 449 ISMGSGKITTIRAKAGHQLFNDPIEMANGNNQPSSEPLKINDGEGYGDIVFANG- 507
QY 418 VMTLILIMWRKLRNPNTYSSLGIVVWSLVSRWGIEMPAIIAR-----SISILSDAG 472
Db 508 -----NSTLYQNV-----IEQGRIVLRERAKLSVNSLSQTG 539
QY 473 --LGMAMFSLGLFWALQPLIACGNKLAALANGVFPVAGPAVMAASTAVGLRGVLLHIA 530
Db 540 GSYMEAGSTLDVFTQPP-----POQPPAANQLITLSN--LHLS 576
QY 531 IVQALPQIV-PFVFAKEYGVHPDIL-STAYGPIITSHG 567
Db 577 LSSILANNAVTPPTNPQDSDHPAVIGSTTAGPVITSG 615

RESULT 13
US-09-612-402B-17
; Sequence 17, Application US/09612402B
; Patent No. 6642023
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCES: 7969-086-999
; CURRENT APPLICATION NUMBER: US/09/612,402B
; CURRENT FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: Patent version 3.0
; SEQ ID NO 17
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-09-612-402B-17

Query Match
4.2%; Score 121; DB 4; Length 505;

Best Local Similarity 22.0%; Pred. No. 0.016;
Matches 125; Conservative 42; Mismatches 152; Indels 248; Gaps 28;

QY 124 LRGMYGASSAGTLMVQVVVLOCIITMYTLMFLFEYRAARALVLDQFP-----DG-- 172
Db 859 LAGMFRAIEBQRLQARLV-----TAPAVQLGDVPLARLAPAVADGAG 900
QY 173 -----AAASIVSFRVDSVVSILAR-----GDVELEAEPDGVAG-----AGAVSSR 212
Db 901 GRDQPAALVWGDGLPLGVVVLAEVVRVGVGAQVAGRLERAAAGVFFQHQHRRHVGVLAAV 960
QY 213 GGDAGRVRVTRKSTSSRSEACSHSHSQTMPRVSNLSGVETIYSSLOSRRNPTPRGSSFN 272
Db 961 VDEVGARPVV-----ELLQDHVTBGAHQGVGALLGVE-----PEVGELG 1001

Best Local Similarity 22.8%; Pred. No. 0.008;
Matches 107; Conservative 56; Mismatches 158; Indels 148; Gaps 25;

QY 36 PDCSG--INRFVALFAVELLSFHF--STNDPFFAMNLRFLAADTLQKV----- 80
Db 50 PLSCFNGLLGSFVLGRGHSITFENIRTSNGAALSN--SAADGLFTIEGKELSPENC 106
QY 81 -AVALL--ALASRGLSSPRALGLDMSITLFSLSLTPVLVVG-----IPLLRGMYGA 130
Db 107 NSLAVLPAATNNGSOTPTTSPNGTIYSKTDL---LLANNEKFSYSLNLSVSGGGTI 163
QY 131 SAGTLMVQVVVLOCIITMYTLMFLFEYRAARALVLDQFPDGAASIV-SFRVDSVWSLA 189
Db 164 IDAKSLTVQGISKLCV-----FQENTAQA-----DGGACQVTSFSAMANEAPI 207
QY 190 RGDVELEAEPDGVAGAGAVSSRGDAGRVRVTRKSTSSRSE-AACSHSHSQTMPRVSN 248
Db 208 A-----FTANVAGVRGGIAAVQDQGG-----VSSSTSTEDPVVFSRNTAVEFDGNVA 257
QY 249 LSGVEIYSLQ-----SSRNPTRGSSFNHADFFNIVGAA-AKG 286
Db 258 RVGG-GIYSYGNVAFNLNGKTLFLNNVASPVYIAAEQPT-SQASNTSNNYDGGGAIFCK 315
QY 287 GCGAAGDEEKG-----ACGGGGGCHSPQOAVAVPAKDKLHMLVWSSS 330
Db 316 NGAQAGSNNSGSVFQGGGVVFFSNVAAKGKG-----AIYAKKLSV-----AN 359
QY 331 ASPV-----SERAHVHVFAG-----GADHADVLAKG-----AQAYDEYGRDDYSSR 372
Db 360 CGPVQFLGNANDGGAIVLGESGELSADYDGIIFDGNLKRKTAKENAADVNGVTVSQA 419
QY 373 TKNGSG-----ADKGG-----PTLSKLSNSTAQ-----LYPKDDGEG 406
Db 420 ISMGSGKITTIRAKAGHQLFNDPIEMANGNNQPSSEPLKINDGEG 468

RESULT 14
US-09-252-991A-20266
; Sequence 20266, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20266
; LENGTH: 1501
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20266

Query Match
4.2%; Score 121; DB 4; Length 1501;

Best Local Similarity 22.0%; Pred. No. 0.041;
Matches 125; Conservative 42; Mismatches 152; Indels 248; Gaps 28;

QY 124 LRGMYGASSAGTLMVQVVVLOCIITMYTLMFLFEYRAARALVLDQFP-----DG-- 172
Db 859 LAGMFRAIEBQRLQARLV-----TAPAVQLGDVPLARLAPAVADGAG 900
QY 173 -----AAASIVSFRVDSVVSILAR-----GDVELEAEPDGVAG-----AGAVSSR 212
Db 901 GRDQPAALVWGDGLPLGVVVLAEVVRVGVGAQVAGRLERAAAGVFFQHQHRRHVGVLAAV 960
QY 213 GGDAGRVRVTRKSTSSRSEACSHSHSQTMPRVSNLSGVETIYSSLOSRRNPTPRGSSFN 272
Db 961 VDEVGARPVV-----ELLQDHVTBGAHQGVGALLGVE-----PEVGELG 1001
```

```

QY 273 HADFFNIYGAAGKGG-----GAAGDEKAGCGGGGGH-----SP 308
Db 1002 H---LGIVRGDGDGLGALVAHLGEBVRGPRRLRHVGCPCDDVAGVVPVGRRLRHVGLLAP 1058
QY 309 QPQA-----VAVPAKRLKDLMLVWSSSASPVSERAAVHVFGAGG-ADHADVLAKGAQAYDE 363
Db 1059 DLRAGRQVAVP-----VVEAQADPADQR---QVAGAGGVADH-----R 1094
QY 364 YGRD-----DYSSRTKNGSGAGKGGPTLSKLGNSSTAQLYPKDDGEGRAAAVAMPASPVM 419
Db 1095 HGRDRREADHPVRTV-GLGGVDVG-----RGDDLGLVLPACTDEAAQAALA----- 1139
QY 420 TRLLIMVWRKLIENPTYSSLIGVWVSLVSYRWGTEMPAIIARSISILSDAGLWAMFS 479
Db 1140 -----LVALRLG-----GVLDLDAGPGVHRVA 1160
QY 480 LGLEMALQPRI-----IACGNKLAAL-AMGVRFVAGPVMMAAASI-----AVGL 522
Db 1161 V-----LRPRLAPELQRLAHLRVLQAVGAVDVFGIAGPARAAARLVVGQVRTGARVVGL 1215
QY 523 RG-----VLLHI-----AIQVQALPGQIVP-----FVF 545
Db 1216 LGFFGDQAVLHVDLPGAGAGVHPVGGAHDLVVLPAIPVGTTPGAALVAGLAIVGEFAL 1275
QY 546 -----AKEYGVHPDILSTAYGEITS 565
Db 1276 LLEELQAVEQVTHLDLITRSVPGVTT 1302

RESULT 15
US-09-252-991A-26256
; Sequence 26256, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26256
; LENGTH: 542
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26256

Query Match 4.1%; Score 119.5; DB 4; Length 542;
Best Local Similarity 23.3%; Pred. No. 0.012;
Matches 101; Conservative 40; Mismatches 158; Indels 135; Gaps 20;

QY 192 GDVELEAPDGVAGAGVSRGGDAGKRVTVTKSTSSRSEACSHSHSQTMQPRVSNLS 251
Db 10 GGVDVGGRRPG-----GAHGNRG-----TAPPAHRPAPSAQPAAD-- 45
QY 252 GVEIYSLQSSRNPTPRGSSFNHADFFNIVGAAAKGGGAAGDEKAGCGGGG----- 304
Db 46 -----PRSPRRRAGACRRGRVQAQADPRPGGAAGARRRPAACGACGAGVAGV 96
QY 305 ---GHSFPQQA---VAVPAK-----RKDLMLVWSSSASPVSFSE-----RAAVH- 341
Db 97 RLRGSRQPAAPAAAPALFGRLAGPVAYGPAACRVGWHARPQAQAPAPAMEDPARAGGHH 156
QY 342 ---VFGAGGADHADVLAQAYDEYGRDDYSRTKNG--SGGADKGGPTLSKLGNSSTA 396
Db 157 LRALAPATAAQARALRAGG-----LSARLQHSRLAAGAAQGGP-----VGAGRLA 202
QY 397 QLYPKDDGEGRAAAVAMP-----PASVMTRLIMVWRKLIENPT 437

```

```

Db 203 R-----PRTVAVAMPARAGGPARARGLDGTLRTGPGACRTGPDVIAAGTCRALRPPV 254
QY 438 YSSLIGVWVHSLVSYRWGTEMPAIIARSISILSDA-----GLGNAMFSLGLF 483
Db 255 FRSAFS--QGSPLRLWALSPREAHMNFSELIOAVRRDPSSVVVPASWQGRATFG-GLV 311
QY 484 MAL-----QPRIIACGNKLAALAMGVRFVAGPVMMAAASIAGVLRGVLLHIAIVQALPQG 539
Db 312 VALAYEAMLAIVVEAGRPLRSI--GVSFVGFLAPEQOPASFSARLLREGKAVSQVQVEVRQG 369
QY 540 --IVPFFVFAKEYGV 551
Db 370 EQVTVLQQA-SFGV 382

```

Search completed: March 3, 2004, 09:09:47  
Job time : 26 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2004, 09:08:49 ; Search time 53 Seconds  
(without alignments)

2282.846 Million cell updates/sec

Title: US-10-030-884-14

Perfect score: 2910  
Sequence: 1 MITALDLXHLVAVPLVYA.....DLSTAYGPIITSHGFIHCS 573

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCT05\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	129.5	4.5	1402	9	US-09-712-363-166
2	129	4.4	646	9	US-09-964-899-53
3	126	4.3	277	15	US-10-032-037B-25
4	126	4.3	277	15	US-10-029-988B-25
5	126	4.3	277	15	US-10-032-423A-25
6	125	4.3	462	10	US-09-919-039-324
7	124.5	4.3	623	13	US-10-108-605-125
8	124.5	4.3	623	13	US-10-108-605-129
9	122.5	4.2	982	9	US-09-841-132-176
10	122.5	4.2	1006	9	US-09-841-132-190
11	122	4.2	277	15	US-10-032-037B-203
12	122	4.2	277	15	US-10-029-988B-203
13	122	4.2	277	15	US-10-032-423A-203
14	122	4.2	323	15	US-10-374-780A-1400
15	121.5	4.2	585	9	US-09-738-626-6032

16	118	4.1	1176	14	US-10-156-761-13466
17	117	4.0	249	10	US-09-880-748-749
18	116	4.0	249	10	US-09-880-748-2033
19	116	4.0	251	10	US-09-880-748-1047
20	115.5	4.0	246	10	US-09-880-748-928
21	115	4.0	249	10	US-09-880-748-772
22	115	4.0	249	10	US-09-880-748-833
23	115	4.0	251	10	US-09-880-748-1053
24	115	4.0	251	10	US-09-880-748-1805
25	115	4.0	266	15	US-10-032-037B-204
26	115	4.0	266	15	US-10-029-988B-204
27	115	4.0	266	15	US-10-032-423A-204
28	115	4.0	892	9	US-09-887-828A-2
29	114.5	3.9	248	10	US-09-880-748-1643
30	114.5	3.9	248	10	US-09-880-748-1681
31	114.5	3.9	423	14	US-10-100-679-95
32	114.5	3.9	471	14	US-10-100-679-96
33	114.5	3.9	541	14	US-10-100-679-97
34	114.5	3.9	582	14	US-10-100-679-81
35	114.5	3.9	1482	10	US-09-940-316B-21
36	114.5	3.9	1488	10	US-09-940-316B-17
37	114	3.9	249	10	US-09-880-748-629
38	114	3.9	251	10	US-09-880-748-1060
39	114	3.9	4471	14	US-10-205-032-10
40	113.5	3.9	246	10	US-09-880-748-2068
41	113.5	3.9	248	10	US-09-880-748-1178
42	113.5	3.9	248	10	US-09-880-748-1306
43	113.5	3.9	248	10	US-09-880-748-1785
44	113.5	3.9	249	10	US-09-880-748-632
45	113.5	3.9	249	10	US-09-880-748-673

#### ALIGNMENTS

#### RESULT 1

US-09-712-363-166  
; Sequence 166, Application US/09712363  
; Patent No. US20020164588A1  
; GENERAL INFORMATION:  
; APPLICANT: Eisenberg, David  
; APPLICANT: Rotstein, Sergio H.  
; APPLICANT: Marcotte, Edward M.  
; TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND  
; TITLE OF INVENTION: INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS  
; FILE REFERENCE: 07419-032001  
; CURRENT APPLICATION NUMBER: US/09/712.363  
; CURRENT FILING DATE: 2000-11-13  
; PRIOR APPLICATION NUMBER: PCT/US00/02246  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: 60/179,531  
; PRIOR FILING DATE: 2000-02-01  
; PRIOR APPLICATION NUMBER: 60/117,844  
; PRIOR FILING DATE: 1999-01-29  
; PRIOR APPLICATION NUMBER: 60/118,206,  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: 60/126,593  
; PRIOR FILING DATE: 1999-03-26  
; PRIOR APPLICATION NUMBER: 60/134,093  
; PRIOR FILING DATE: 1999-05-14  
; PRIOR APPLICATION NUMBER: 60/134,092  
; PRIOR FILING DATE: 1999-05-14  
; PRIOR APPLICATION NUMBER: 60/165,124  
; PRIOR FILING DATE: 1999-11-12  
; PRIOR APPLICATION NUMBER: 60/165,086  
; PRIOR FILING DATE: 1999-11-12  
; NUMBER OF SEQ ID NOS: 292  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 166  
; LENGTH: 1402  
; TYPE: PRT  
; ORGANISM: Mycobacterium tuberculosis  
US-09-712-363-166

Sequence 13466, A  
Sequence 749, App  
Sequence 2033, Ap  
Sequence 1047, Ap  
Sequence 928, App  
Sequence 833, App  
Sequence 1053, Ap  
Sequence 1805, Ap  
Sequence 204, App  
Sequence 204, App  
Sequence 2, Appl  
Sequence 1643, Ap  
Sequence 1681, Ap  
Sequence 95, Appl  
Sequence 96, Appl  
Sequence 97, Appl  
Sequence 81, Appl  
Sequence 21, Appl  
Sequence 17, Appl  
Sequence 629, App  
Sequence 1060, Ap  
Sequence 10, Appl  
Sequence 2068, Ap  
Sequence 1178, Ap  
Sequence 1306, Ap  
Sequence 1755, Ap  
Sequence 632, App  
Sequence 673, App



```

QY 346 GGADHADVLAKGAQVDEYGRDDY--SSRTKNGSGGADKGGFTLSKLSNSTAQLYPKDD 403
DB 216 SSGNTASLTITGAQAED---ADYYCNSRDSGNHVVFGGTTKLTVLGAAAEQKLISEED 272
QY 404 GEGRA 408
DB 273 LINGAA 277

```

## RESULT 4

```

US-10-029-988B-25
; Sequence 25, Application US/10029988B
; Publication No. US20040001839A1
; GENERAL INFORMATION:
; APPLICANT: Bio-Technology General Corp.
; TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
; FILE OF INVENTION: MOLETTIES, ANTIBODIES TO SUCH EPITOPES, AND USES THEREOF
; FILE REFERENCE: 10793/45
; CURRENT APPLICATION NUMBER: US/10/029,988B
; CURRENT FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/258,948
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 25
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-029-988B-25

```

```

Query Match 4.3%; Score 126; DB 15; Length 277;
Best Local Similarity 22.3%; Pred. No. 0.0063;
Matches 68; Conservative 36; Mismatches 125; Indels 76; Gaps 9;

```

```

QY 152 MLFLFEYRAARALVLDQFPDGAASIV-----SFRVDSVVSRLAR 191
DB 1 MKYLLPTAAAGLLLLAAQAPAMAEVOLVESGGVVRPGGSLRLSCAASGFTDDYGMWVR 60
QY 192 GDVELEAEPDGVAGAGAVSSRGDAGRVRVTVRKSTSRSEACSHSHSQTMQPRVSNLS 251
DB 61 -----QAPKGLEWVSGINWNGGSTGYADSVKGRFTISRDN-----KNSLYLQNSLRAE 111
QY 252 GVEIYLSQSSRNTPRGSSFNHADFFNIVGAAAGGGAAGDEKAGCGGGGSHSPQP 311
DB 112 DTAVVYICARMRAPVIWQ-----GTLVTVSRGGGSGG-----GGSGGGSSSELTQDP 159
QY 312 AVAV-----PAKXKDLHMLVWSSASPVSEAAVHVFGA 345
DB 160 AVSVALGQTVRITCQDSLSRYASYWYQKPGQAPVLIYGNRPSPGIPDRFS-----GS 215
QY 346 GGADHADVLAKGAQVDEYGRDDY--SSRTKNGSGGADKGGFTLSKLSNSTAQLYPKDD 403
DB 216 SSGNTASLTITGAQAED---ADYYCNSRDSGNHVVFGGTTKLTVLGAAAEQKLISEED 272
QY 404 GEGRA 408
DB 273 LINGAA 277

```

## RESULT 5

```

US-10-032-423A-25
; Sequence 25, Application US/10032423A
; Publication No. US20040002450A1
; GENERAL INFORMATION:
; APPLICANT: Bio-Technology General Corp.
; TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
; FILE OF INVENTION: MOLETTIES, ANTIBODIES TO SUCH EPITOPES, AND USES THEREOF
; FILE REFERENCE: 10793/45
; CURRENT APPLICATION NUMBER: US/10/032,423A
; CURRENT FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/258,948
; PRIOR FILING DATE: 12/29/2000

```

```

; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 25
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-032-423A-25

```

```

Query Match 4.3%; Score 126; DB 15; Length 277;
Best Local Similarity 22.3%; Pred. No. 0.0063;
Matches 68; Conservative 36; Mismatches 125; Indels 76; Gaps 9;
QY 152 MLFLFEYRAARALVLDQFPDGAASIV-----SFRVDSVVSRLAR 191
DB 1 MKYLLPTAAAGLLLLAAQAPAMAEVOLVESGGVVRPGGSLRLSCAASGFTDDYGMWVR 60
QY 192 GDVELEAEPDGVAGAGAVSSRGDAGRVRVTVRKSTSRSEACSHSHSQTMQPRVSNLS 251
DB 61 -----QAPKGLEWVSGINWNGGSTGYADSVKGRFTISRDN-----KNSLYLQNSLRAE 111
QY 252 GVEIYLSQSSRNTPRGSSFNHADFFNIVGAAAGGGAAGDEKAGCGGGGSHSPQP 311
DB 112 DTAVVYICARMRAPVIWQ-----GTLVTVSRGGGSGG-----GGSGGGSSSELTQDP 159
QY 312 AVAV-----PAKXKDLHMLVWSSASPVSEAAVHVFGA 345
DB 160 AVSVALGQTVRITCQDSLSRYASYWYQKPGQAPVLIYGNRPSPGIPDRFS-----GS 215
QY 346 GGADHADVLAKGAQVDEYGRDDY--SSRTKNGSGGADKGGFTLSKLSNSTAQLYPKDD 403
DB 216 SSGNTASLTITGAQAED---ADYYCNSRDSGNHVVFGGTTKLTVLGAAAEQKLISEED 272
QY 404 GEGRA 408
DB 273 LINGAA 277

```

## RESULT 6

```

US-09-919-039-324
; Sequence 324, Application US/09919039
; Publication No. US20030108871A1
; GENERAL INFORMATION:
; APPLICANT: Kaet, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 324
; LENGTH: 462
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030108871A1 1813444CD1
US-09-919-039-324

```

```

Query Match 4.3%; Score 125; DB 10; Length 462;
Best Local Similarity 27.1%; Pred. No. 0.016;
Matches 56; Conservative 18; Mismatches 75; Indels 58; Gaps 10;
QY 208 AVSSRGGDAGRVRVTVRKSTSRSEACSHSHSQTMQPRVSNLSGVETISLQSRNTPR 267
DB 105 APSSTSGSYG-----SSQSSSYGQPSGYSQPSYGGQ--QSYGQSQSYNP-PQ 153
QY 268 GGSFNHADFFNIVGAAAGGGAAGDEKAGCGGGGSHSPQPQAVAPAKRDLMLVW 327
DB 154 G--YGGQNYN-----SSSGGGG-----GGGGGNYGQD----- 181
QY 328 SSSASPVSEAAVHVFGAGGADHADVLAKGAQVDEYGRDDYSSRTKNGSGGADKGGFTL 387

```

Db 182 SNSSG-----GSGGGYGNQDSGGGGGGYGGQDRGRGCGSGGGGGGGG- 230  
Qy 389 SKLGSNSTAQLY-PKDDGEGRAAAVAM 413  
Db 231 ---GYNRSSGVEPRGGGGRGGM 254

## RESULT 7

US-10-108-605-125  
; Sequence 125, Application US/10108605  
; Publication No. US20020160934A1  
; GENERAL INFORMATION:  
; APPLICANT: Broadus, Julie  
; APPLICANT: Stam, Lynn  
; APPLICANT: Bachmann, Jane  
; APPLICANT: Kamdar, Kim  
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE  
; FILE REFERENCE: 31133B  
; CURRENT APPLICATION NUMBER: US/10/108,605  
; CURRENT FILING DATE: 2002-03-27  
; PRIOR APPLICATION NUMBER: US 09/761,142  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/176,418  
; PRIOR FILING DATE: 2000-01-14  
; NUMBER OF SEQ ID NOS: 361  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 125  
; LENGTH: 623  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-10-108-605-125

Query Match 4.3%; Score 124.5; DB 13; Length 623;  
Best Local Similarity 20.6%; Pred. No. 0.027;  
Matches 80; Conservative 55; Mismatches 134; Indels 119; Gaps 15;  
Qy 161 ARAVLDPDPDGA--AASIVSRVDSVVS LARG-----DVELEAEPDGV 203  
Db 196 AASLGLGYFNDMAFFVGDANAYTDSVNFSSGYNTSNTHTDRINNSTPPQQSQQPTV 255  
Qy 204 AGAGAVSSRGDGRVVRVTRKSTSRBAACSHSHSQTMPQPRVNSLGVETIYLSQSRN 263  
Db 256 NGSGSASSNNNS--MLPPAVQQSNENNNTSSNTNNSSNN---NNNSGGSNNNAGSNN 311  
Qy 264 PTPRGSSFNHADFFNIVGAAA-----KGGGGAAGDEEKGACGGGGGHSPOQAVAVPA 317  
Db 312 -----NNNNNNINFMAAAIAFQHLKKEPPT-----QNGNIGYGGGNSQNDPTDLS 361  
Qy 318 KRKDLHMLVW--SSSASPVSERAAVHVFAGGAD---HADVLAKGAQAYDEYGRDDYSSR 372  
Db 362 YGLPAHLAAYGGGSGGPTGGRSS---GGGDES DYHSTISAQDHQSQSSG-----410  
Qy 373 TKNGSGADKGGPTLSKLSNSTAQLYPKDDGGR-----IMVWRKLI-----RN 434  
Db 411 -GNGSGGASGGTSGNSGYLDSSEFVGSYAGNRFRHDGYPPFTPYDAQSFQSMGPQPT 469  
Qy 408 -----AAAVAMPASVMTLIL-----VSRYWGI 455  
Db 470 AMDQWGAHAHQHPAAVYMTSLGLDKLGGYTTQGGVPCFTGSGPIQLWQFLLELLDKT 529  
Qy 435 PNTYSSLIQVWVSL-----VSRYWGI 455  
Db 530 CQSFISWTGDEWEPKLTDPDEVARRNGI 557

## RESULT 8

US-10-108-605-129  
; Sequence 129, Application US/10108605  
; Publication No. US20020160934A1  
; GENERAL INFORMATION:  
; APPLICANT: Broadus, Julie

; APPLICANT: Stam, Lynn  
; APPLICANT: Bachmann, Jane  
; APPLICANT: Kamdar, Kim  
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE  
; FILE REFERENCE: 31133B  
; CURRENT APPLICATION NUMBER: US/10/108,605  
; CURRENT FILING DATE: 2002-03-27  
; PRIOR APPLICATION NUMBER: US 09/761,142  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/176,418  
; PRIOR FILING DATE: 2000-01-14  
; NUMBER OF SEQ ID NOS: 361  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 129  
; LENGTH: 623  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-10-108-605-129

Query Match 4.3%; Score 124.5; DB 13; Length 623;  
Best Local Similarity 20.6%; Pred. No. 0.027;  
Matches 80; Conservative 55; Mismatches 134; Indels 119; Gaps 15;  
Qy 161 ARAVLDPDPDGA--AASIVSRVDSVVS LARG-----DVELEAEPDGV 203  
Db 196 AASLGLGYFNDMAFFVGDANAYTDSVNFSSGYNTSNTHTDRINNSTPPQQSQQPTV 255  
Qy 204 AGAGAVSSRGDGRVVRVTRKSTSRBAACSHSHSQTMPQPRVNSLGVETIYLSQSRN 263  
Db 256 NGSGSASSNNNS--MLPPAVQQSNENNNTSSNTNNSSNN---NNNSGGSNNNAGSNN 311  
Qy 264 PTPRGSSFNHADFFNIVGAAA-----KGGGGAAGDEEKGACGGGGGHSPOQAVAVPA 317  
Db 312 -----NNNNNNINFMAAAIAFQHLKKEPPT-----QNGNIGYGGGNSQNDPTDLS 361  
Qy 318 KRKDLHMLVW--SSSASPVSERAAVHVFAGGAD---HADVLAKGAQAYDEYGRDDYSSR 372  
Db 362 YGLPAHLAAYGGGSGGPTGGRSS---GGGDES DYHSTISAQDHQSQSSG-----410  
Qy 373 TKNGSGADKGGPTLSKLSNSTAQLYPKDDGGR-----IMVWRKLI-----RN 434  
Db 411 -GNGSGGASGGTSGNSGYLDSSEFVGSYAGNRFRHDGYPPFTPYDAQSFQSMGPQPT 469  
Qy 408 -----AAAVAMPASVMTLIL-----VSRYWGI 455  
Db 470 AMDQWGAHAHQHPAAVYMTSLGLDKLGGYTTQGGVPCFTGSGPIQLWQFLLELLDKT 529  
Qy 435 PNTYSSLIQVWVSL-----VSRYWGI 455  
Db 530 CQSFISWTGDEWEPKLTDPDEVARRNGI 557

## RESULT 9

US-09-841-132-176  
; Sequence 176, Application US/09841132  
; Patent No. US20020061848A1  
; GENERAL INFORMATION:  
; APPLICANT: Skeiky, Yasar A.W.  
; APPLICANT: Probst, Peter  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
; FILE REFERENCE: 210121.469CB  
; CURRENT APPLICATION NUMBER: US/09/841,132  
; CURRENT FILING DATE: 2001-04-23  
; NUMBER OF SEQ ID NOS: 599  
; SOFTWARE: FastSeq for Windows Version 3.0/4.0  
; SEQ ID NO 176  
; LENGTH: 582  
; TYPE: PRT  
; ORGANISM: Chlamydia  
; FEATURE:

NAME/KEY: VARIANT  
LOCATION: (1)...(982)  
OTHER INFORMATION: Xaa = Any Amino Acid  
US-09-841-132-176

Query Match  
Best Local Similarity 4.2%; Score 122.5; DB 9; Length 982;  
Matches 137; Conservative 71; Mismatches 222; Indels 209; Gaps 32;

QY 36 PDQCSG--INRFVAFVPLLSFHF--STNDPFFAMNLFRLAADTLQKV-----80  
DB 48 PLSCFGNLSGFTVLGRGHSITFNIRTSNGAALSN---SAADGLFTIEGFKELSFNSC 104  
QY 81 -AVLALL--ALASRGLSSPRALGLDWSITLFSLSSTLPTNLVWG-----IPLRGMVGA 130  
DB 105 NSLLAVLPAATTKNGSQTPTTSTPSNGTIYSKTDL---LLNNEKFSFYSNLVSDGGA 161  
QY 131 SSAGTLMVQVVLQCIWYTLMLFLFEYRAARALVLDQPDGAAASIV-SFRVSDVWSL 189  
DB 162 IDAKSLTVQIGSKLCV-----FOENTAQA-----DGAACQVTSFMANEAPI 205  
QY 190 ARGVLELEAPDGVAGAGVSSRGDAGRVRVTRKSTSRSE-AACSHSHSQTWQPRVS 248  
DB 206 A-----FVANVAGVGGGIAAQQDQGG-----VSSSTSTEDPVVSFRNTAVEFDGVA 255  
QY 249 NLGSGVEIYSLQ-----SSRNPPTRGSSFNHADFNIVGAAAKG 286  
DB 256 RVGG-GIYSYGNVAFNLNGKTLFLNNVASPVYIAAKQPT-SGAASNTSNNGYGGGAI FCK 313  
QY 287 GGGAGDEEKG-----ACGGGGGHSPOQAVAPAKKDLHMLVWSSS 330  
DB 314 NGAAGSNNSSVSFDGEGVVFSSNVAAGKGG-----AIYAKKLSV-----AN 357  
QY 331 ASPV-----SRAAHVHVFAG-----GADHADVLAKG-----AAQVDEYGRDDYSR 372  
DB 358 CGPQVFLRNTANDGGAIIYLGESGELSADYDGIIFDGNLKRKTAKENAADVNGVTSSOA 417  
QY 373 TKNGSGG-----ADKGG-----PTLSKLGSNSTAQ-----LYPKDDEGRAAAVAMPAS 417  
DB 418 ISMGSGGKITTLRAKAGHQILFNDPIEMANGNQPACSSKLLKINDGEGYTGDI VFANGS 477  
QY 418 VMTRLILIMVWRKLIERNPTYSSLIGVWVSLVSRWGIEMPAIAR-----SISILSDAG 472  
DB 478 -----STIQVNTIEQGRIVLRREKAKLSVNSLSQTG 508  
QY 473 --LGMAMFSLGLFMALQPRIIACGNKLAIAIANGVRFVAGPAAVMAAASIAVGLRGVLLHIA 530  
DB 509 GSLYMEAGSTLDFVTQP-----PQPPAANQLITLSN--LHLS 545  
QY 531 IVQAAALPQGIIV-PFVFAKEYGVHPDIL-STAYGPITSHG 567  
DB 546 LSSLIANNVNTNPTNPPAQDSHPAVIGSTTAGSVTISG 584

RESULT 10  
US-09-841-132-190  
Sequence 190, Application US/09841132  
Patent No. US20020061848A1  
GENERAL INFORMATION:  
APPLICANT: Bhatia, Ajay  
APPLICANT: Skeiky, Yasir A.W.  
TITLE OF INVENTION: PROBST, Peter  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION  
FILE REFERENCE: 210121.469C8  
CURRENT APPLICATION NUMBER: US/09/841.132  
CURRENT FILING DATE: 2001-04-23  
NUMBER OF SEQ ID NOS: 599  
SOFTWARE: FastSeq for Windows Version 3.0/4.0  
SEQ ID NO 190  
LENGTH: 1006  
TYPE: PR  
ORGANISM: Chlamydia

US-09-841-132-190

Query Match  
Best Local Similarity 4.2%; Score 122.5; DB 9; Length 1006;  
Matches 137; Conservative 71; Mismatches 222; Indels 209; Gaps 32;

QY 36 PDQCSG--INRFVAFVPLLSFHF--STNDPFFAMNLFRLAADTLQKV-----80  
DB 72 PLSCFGNLSGFTVLGRGHSITFNIRTSNGAALSN---SAADGLFTIEGFKELSFNSC 128  
QY 81 -AVLALL--ALASRGLSSPRALGLDWSITLFSLSSTLPTNLVWG-----IPLRGMVGA 130  
DB 129 NSLLAVLPAATTKNGSQTPTTSTPSNGTIYSKTDL---LLNNEKFSFYSNLVSDGGA 185  
QY 131 SSAGTLMVQVVLQCIWYTLMLFLFEYRAARALVLDQPDGAAASIV-SFRVSDVWSL 189  
DB 186 IDAKSLTVQIGSKLCV-----FOENTAQA-----DGAACQVTSFMANEAPI 229  
QY 190 ARGVLELEAPDGVAGAGVSSRGDAGRVRVTRKSTSRSE-AACSHSHSQTWQPRVS 248  
DB 230 A-----FVANVAGVGGGIAAQQDQGG-----VSSSTSTEDPVVSFRNTAVEFDGVA 279  
QY 249 NLGSGVEIYSLQ-----SSRNPPTRGSSFNHADFNIVGAAAKG 286  
DB 280 RVGG-GIYSYGNVAFNLNGKTLFLNNVASPVYIAAKQPT-SGAASNTSNNGYGGGAI FCK 337  
QY 287 GGGAGDEEKG-----ACGGGGGHSPOQAVAPAKKDLHMLVWSSS 330  
DB 338 NGAAGSNNSSVSFDGEGVVFSSNVAAGKGG-----AIYAKKLSV-----AN 381  
QY 331 ASPV-----SRAAHVHVFAG-----GADHADVLAKG-----AAQVDEYGRDDYSR 372  
DB 382 CGPQVFLRNTANDGGAIIYLGESGELSADYDGIIFDGNLKRKTAKENAADVNGVTSSOA 441  
QY 373 TKNGSGG-----ADKGG-----PTLSKLGSNSTAQ-----LYPKDDEGRAAAVAMPAS 417  
DB 442 ISMGSGGKITTLRAKAGHQILFNDPIEMANGNQPACSSKLLKINDGEGYTGDI VFANGS 501  
QY 418 VMTRLILIMVWRKLIERNPTYSSLIGVWVSLVSRWGIEMPAIAR-----SISILSDAG 472  
DB 502 -----STIQVNTIEQGRIVLRREKAKLSVNSLSQTG 532  
QY 473 --LGMAMFSLGLFMALQPRIIACGNKLAIAIANGVRFVAGPAAVMAAASIAVGLRGVLLHIA 530  
DB 533 GSLYMEAGSTLDFVTQP-----PQPPAANQLITLSN--LHLS 569  
QY 531 IVQAAALPQGIIV-PFVFAKEYGVHPDIL-STAYGPITSHG 567  
DB 570 LSSLIANNVNTNPTNPPAQDSHPAVIGSTTAGSVTISG 608

RESULT 11  
US-10-032-037B-203  
Sequence 203, Application US/10032037B  
Publication No. US20040001822A1  
GENERAL INFORMATION:  
APPLICANT: Bio-Technology General Corp.  
TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED  
TITLE OF INVENTION: MOIETIES, ANTIBODIES TO SUCH EPITOPES, AND USES THEREOF  
FILE REFERENCE: 10793/44  
CURRENT APPLICATION NUMBER: US/10/032,037B  
CURRENT FILING DATE: 2001-12-31  
PRIOR APPLICATION NUMBER: 60/258,948  
PRIOR FILING DATE: 2000-12-29  
NUMBER OF SEQ ID NOS: 204  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 203  
LENGTH: 277  
TYPE: PR  
ORGANISM: Homo sapiens  
US-10-032-037B-203

Query Match  
4.2%; Score 122; DB 15; Length 277;



Best Local Similarity 21.8%; Pred. No. 0.014;  
Matches 67; Conservative 40; Mismatches 120; Indels 80; Gaps 9;  
QY 152 MLFLFYRAARALVLDQFPDGAASIV-----SPRVSDVVSLAR 191  
Db 1 MKYLLPTAAAGLLLLAAQPAVAEVLVESGGVVRPGGSLRLSCAASGFTFDDYGNMSVR 60  
QY 192 GDVELEAPDGVAGAGAVSSRGDAGRVVTRVTKSTSSRSEAAACSHSHSQTMQPRVSNLS 251  
Db 61 -----QAPCKGLEWVSGINWNGSGTGADSVKGRFTISRDN-----KNSL 101  
QY 252 GVEIYSLQSSRNPTPRGSSFNHADFF--NIVGAAKGGGGAAGDEBKACGGGGHSPQ 309  
Db 102 YLQMSLSRAEDTAVYICARLTHPYFWGQGLTVTVSRGGGGSGG---GGSGGGGSSSELTQ 157  
QY 310 PQAVAV-----PAKRDLEMLVWSSASPVSEAAAHVF 343  
Db 158 DPAVSVALGQTVITCQDLSRSYASWYQKPGQAPVLIYKNNRPSGIPDRFS--- 213  
QY 344 GAGGADHADVLAKGAQAYDEYGRDYY--SSRTKNGSGGADKGGPTLSKLSNSTAQLYPK 401  
Db 214 GSSSGNTASLTITGAQAED---ADYYCNSRDSGNHVVFGGQTKLTVLGAAGAEQKLISE 270  
QY 402 DDGEGR 408  
Db 271 EDLNGAA 277

RESULT 12  
US-10-029-988B-203  
; Sequence 203, Application US/10029988B  
; Publication No. US20040001839A1  
; GENERAL INFORMATION:  
; APPLICANT: Bio-Technology General Corp.  
; TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED  
; FILE REFERENCE: 10793/46  
; CURRENT APPLICATION NUMBER: US/10/029,988B  
; PRIOR FILING DATE: 2001-12-31  
; PRIOR FILING DATE: 2000-12-29  
; NUMBER OF SEQ ID NOS: 204  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 203  
; LENGTH: 277  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-029-988B-203

Query Match 4.2%; Score 122; DB 15; Length 277;  
Best Local Similarity 21.8%; Pred. No. 0.014;  
Matches 67; Conservative 40; Mismatches 120; Indels 80; Gaps 9;  
QY 152 MLFLFYRAARALVLDQFPDGAASIV-----SPRVSDVVSLAR 191  
Db 1 MKYLLPTAAAGLLLLAAQPAVAEVLVESGGVVRPGGSLRLSCAASGFTFDDYGNMSVR 60  
QY 192 GDVELEAPDGVAGAGAVSSRGDAGRVVTRVTKSTSSRSEAAACSHSHSQTMQPRVSNLS 251  
Db 61 -----QAPCKGLEWVSGINWNGSGTGADSVKGRFTISRDN-----KNSL 101  
QY 252 GVEIYSLQSSRNPTPRGSSFNHADFF--NIVGAAKGGGGAAGDEBKACGGGGHSPQ 309  
Db 102 YLQMSLSRAEDTAVYICARLTHPYFWGQGLTVTVSRGGGGSGG---GGSGGGGSSSELTQ 157  
QY 310 PQAVAV-----PAKRDLEMLVWSSASPVSEAAAHVF 343  
Db 158 DPAVSVALGQTVITCQDLSRSYASWYQKPGQAPVLIYKNNRPSGIPDRFS--- 213  
QY 344 GAGGADHADVLAKGAQAYDEYGRDYY--SSRTKNGSGGADKGGPTLSKLSNSTAQLYPK 401  
Db 214 GSSSGNTASLTITGAQAED---ADYYCNSRDSGNHVVFGGQTKLTVLGAAGAEQKLISE 270

QY 402 DDGEGR 408  
Db 271 EDLNGAA 277  
RESULT 13  
US-10-032-423A-203  
; Sequence 203, Application US/10032423A  
; Publication No. US20040002450A1  
; GENERAL INFORMATION:  
; APPLICANT: Bio-Technology General Corp.  
; TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED  
; FILE REFERENCE: 10793/45  
; CURRENT APPLICATION NUMBER: US/10/032,423A  
; PRIOR FILING DATE: 2001-12-31  
; PRIOR FILING DATE: 2000-12-29  
; NUMBER OF SEQ ID NOS: 204  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 203  
; LENGTH: 277  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-032-423A-203

Query Match 4.2%; Score 122; DB 15; Length 277;  
Best Local Similarity 21.8%; Pred. No. 0.014;  
Matches 67; Conservative 40; Mismatches 120; Indels 80; Gaps 9;  
QY 152 MLFLFYRAARALVLDQFPDGAASIV-----SPRVSDVVSLAR 191  
Db 1 MKYLLPTAAAGLLLLAAQPAVAEVLVESGGVVRPGGSLRLSCAASGFTFDDYGNMSVR 60  
QY 192 GDVELEAPDGVAGAGAVSSRGDAGRVVTRVTKSTSSRSEAAACSHSHSQTMQPRVSNLS 251  
Db 61 -----QAPCKGLEWVSGINWNGSGTGADSVKGRFTISRDN-----KNSL 101  
QY 252 GVEIYSLQSSRNPTPRGSSFNHADFF--NIVGAAKGGGGAAGDEBKACGGGGHSPQ 309  
Db 102 YLQMSLSRAEDTAVYICARLTHPYFWGQGLTVTVSRGGGGSGG---GGSGGGGSSSELTQ 157  
QY 310 PQAVAV-----PAKRDLEMLVWSSASPVSEAAAHVF 343  
Db 158 DPAVSVALGQTVITCQDLSRSYASWYQKPGQAPVLIYKNNRPSGIPDRFS--- 213  
QY 344 GAGGADHADVLAKGAQAYDEYGRDYY--SSRTKNGSGGADKGGPTLSKLSNSTAQLYPK 401  
Db 214 GSSSGNTASLTITGAQAED---ADYYCNSRDSGNHVVFGGQTKLTVLGAAGAEQKLISE 270  
QY 402 DDGEGR 408  
Db 271 EDLNGAA 277

RESULT 14  
US-10-374-780A-1400  
; Sequence 1400, Application US/10374780A  
; Publication No. US20040019927A1  
; GENERAL INFORMATION:  
; APPLICANT: Sherman, Bradley K  
; APPLICANT: Riechmann, Jose Luis  
; APPLICANT: Jiang, Cai-Zhong  
; APPLICANT: Heard, Jacqueline E  
; APPLICANT: Haake, Volker  
; APPLICANT: Creelman, Robert A  
; APPLICANT: Ratcliffe, Oliver  
; APPLICANT: Adam, Luc J  
; APPLICANT: Reuber, T. Lynne  
; APPLICANT: Keddle, James  
; APPLICANT: Broun, Pierre E  
; APPLICANT: Pilgrim, Marsha L  
; APPLICANT: Dubell III, Arnold T

APPLICANT: Pineda, Omaira  
APPLICANT: YU, Guo-Liang  
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS  
FILE REFERENCE: MBI-0047 CIP  
CURRENT APPLICATION NUMBER: US/10/374,780A  
CURRENT FILING DATE: 2003-02-25  
PRIOR APPLICATION NUMBER: 09/837,944  
PRIOR FILING DATE: 2001-04-18  
PRIOR APPLICATION NUMBER: 60/310,847  
PRIOR FILING DATE: 2001-08-09  
PRIOR APPLICATION NUMBER: 09/934,455  
PRIOR FILING DATE: 2001-08-22  
PRIOR APPLICATION NUMBER: 60/336,049  
PRIOR FILING DATE: 2001-11-19  
PRIOR APPLICATION NUMBER: 60/338,692  
PRIOR FILING DATE: 2001-12-11  
PRIOR APPLICATION NUMBER: 10/171,468  
PRIOR FILING DATE: 2002-06-14  
PRIOR APPLICATION NUMBER: 10/225,066  
PRIOR FILING DATE: 2002-08-09  
PRIOR APPLICATION NUMBER: 10/225,067  
PRIOR FILING DATE: 2002-08-09  
PRIOR APPLICATION NUMBER: 10/225,068  
PRIOR FILING DATE: 2002-08-09  
NUMBER OF SEQ ID NOS: 2906  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 1400  
LENGTH: 323  
TYPE: PRT  
ORGANISM: Oryza sativa  
FEATURE:  
OTHER INFORMATION: Orthologous to G1073  
US-10-374-780A-1400

Query Match 4.2%; Score 122; DB 15; Length 323;  
Best Local Similarity 23.0%; Pred. No. 0.018;  
Matches 87; Conservative 37; Mismatches 122; Indels 132; Gaps 18;  
QY 200 PDGVAGAVSSRGDAGRVTVRKTSRSEAAACSHSHSQTMOPR-----VSNLSGVEI 255  
Db 30 PDDTAMAG-MDPGGGGAG-----AGSSR-----YFHLLRPQPSPLSPTSHVKM 76  
QY 256 YLSQSSRNPTPRGSSFNHADFNIVGAAAKGGGGAAGDEE-----KGACGG-G 302  
Db 77 EHSKMSDPKSPVREG-DH-----AGSGSGGVGGDHQPSSSAMVPVEGSGSAGSG 127  
QY 303 GGHSPQPOAVAPAKKDLHMLVWSSSASPVSERAAHVHVGAGGADHADVLAAGQAYD 362  
Db 128 SGGPTRRPRGRPPGSKNPKPPIITVTRDSPNALHSHVLEV--AGADVDCVA----- 178  
QY 363 EYGRDDYSSRTKNG-----SGGADKGGPTLSKLSNSTAQLYPKDDGEGRAAAVAMPASV 418  
Db 179 -----EYARRRGRGVCLSGGAVVNVALLRQPG-----ASPPGSM 213  
QY 419 MTRLILIMVWRKLIARNPTYSLSLGVVWSLVSVYRWGTEMPALIAHSIILSDAGLGNAMF 478  
Db 214 VATL-----RGRFEILSLGTG-----LPPAPPAGAS-----GLTVF 245  
QY 479 SLGLFMALQPRIACGNKLAIAIMGVRFVAGPAPVMAAASIA-----VGLRGVLLHIA-- 530  
Db 246 LSG-----GGQVIGSVVGPVLAAGPVVLAASPANAYVRLPLEGESEEEVYAP 295  
QY 531 -----IVQAALPQG 539  
Db 296 AAGGEAQDQVAGSAGPQG 313

RESULT 15  
US-09-738-626-6032  
Sequence 6032, Application US/09738626  
Publication No. US20020197605A1  
GENERAL INFORMATION:  
APPLICANT: NAKAGAWA, SATOSHI

APPLICANT: MIZOGUCHI, HIROSHI  
APPLICANT: ANDO, SEIKO  
APPLICANT: HAYASHI, MIKIRO  
APPLICANT: OCHIAI, KEIKO  
APPLICANT: YOKOI, HARUHIKO  
APPLICANT: TATEISHI, NAKO  
APPLICANT: SENOH, AKIHIRO  
APPLICANT: IKEDA, MASATO  
APPLICANT: OZAKI, AKIO  
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
FILE REFERENCE: 249-125  
CURRENT APPLICATION NUMBER: US/09/738,626  
CURRENT FILING DATE: 2000-12-18  
PRIOR APPLICATION NUMBER: JP 99/377484  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: JP 00/159162  
PRIOR FILING DATE: 2000-04-07  
PRIOR APPLICATION NUMBER: JP 00/280988  
PRIOR FILING DATE: 2000-08-03  
NUMBER OF SEQ ID NOS: 7059  
SOFTWARE: PatentIn ver. 3.0  
SEQ ID NO 6032  
LENGTH: 585  
TYPE: PRT  
ORGANISM: Corynebacterium glutamicum  
US-09-738-626-6032

Query Match 4.2%; Score 121.5; DB 9; Length 585;  
Best Local Similarity 22.4%; Pred. No. 0.046;  
Matches 82; Conservative 50; Mismatches 147; Indels 87; Gaps 15;  
QY 65 FAMNLRFLAAD--TLQKVAV-----LALLALASRLSSP-RALGLDWSITLFSLSLTPT 116  
Db 209 FAANAKMGASVGTNLRIYMGHMQVMAATSIIKAILDPVQRLAAERAFASFQATFOAD 268  
QY 117 LVWGIP---LLRGMGAS-SAGTLMVVQVVLQIIWYILMLFLFYRAARALVLDQF--- 169  
Db 269 VLTGMPPVSNLMQKMGANGAGEI-----ALGMDIEIAGS 302  
QY 170 -----PDGAAASIVSFVDSVVSRLARGDVLEBAEPDGVAGAVSSRGDAGR 219  
Db 303 GQANSAAGLTPGSAAGGVA-----NAGSIAPDAAVQGAAGQSGVSGFTVTDQL 352  
QY 220 -RVYRKSTSRSEAAACSHSHSQTMOPRVSNLSGVEIYLSQSSRNPTPRGSSFNHADFN 278  
Db 353 DGINIGDMLTSAASAGQSLANGLAMPTSTPNSASGAIPSSMSAASPL-----GAFGS 404  
QY 279 IVCAAAKGG-----GAAGDEEKACGGGGGSHSPQPOAVAPAKKDLHMLVWSSSAS 332  
Db 405 GAGLGAQGGSISSAPGALISSRAAGSAGSVPMGTGGPGA---PGITSDSLNGARTHGAS 461  
QY 333 PVSEAAHVHVGAGGADHADVLAAGQAYDEYGRDDYS---SRTKNGSG---GADKGGP 385  
Db 462 SAGAVAPM-MCGAGGWSGGVVGAGGTGSGSKYARQTGSSVSSSSQSGSLGMVSGSGKP 520  
QY 386 TLSKLG 391  
Db 521 SISNFG 526

Search completed: March 3, 2004, 09:14:08  
Job time : 56 secs